



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121196

TO: Manjunath N Rao
Location: rem/3b11/3c70
Art Unit: 1652
Friday, May 14, 2004
Case Serial Number: 10/016725

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: (571)272-2527

paul.schulwitz@uspto.gov

Search Notes

Examiner Rao,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(571)272-2527

STIC-Biotech/ChemLib

121196

From: Rao, Manjunath N.
Sent: Wednesday, May 05, 2004 10:22 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 10/016,725

From: Manjunath N. Rao
Art Unit 1652, Room ~~10711~~
Mail Box in Room ~~10711~~
Phone: 306-5681

Date: 5-5-04

Please search the following as soon as possible for application with serial number **10/016725**

1. Nucleotides 1-3958 of SEQ ID NO: 15

2. Nucleotides 1-3910 of SEQ ID NO:15

against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao
Art Unit 1652, Room 3B81,
Mail Box in Room 3C70,
Remsen Building, USPTO
400, Dulany St.
Alexandria, VA.
Phone: 571-272-0939

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 5/4
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 23820.8 Seconds
(without alignments)
7114.428 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910
Perfect score: 3910
Sequence: 1 aggaattcattcatttaataa.....cggacacggtcaccgcgttgc 3910

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pla.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3910	100.0	4150	6	AX528729	Sequence
	3910	100.0	4150	9	AF320307	AF320307 Homo sapi
C 3	3880.2	99.2	159351	2	AC018783	AC018783 Homo sapi
	3869.8	99.0	115278	9	AL139819	AL139819 Human DNA
C 5	1779.2	45.5	185644	2	AL391723	AL391723 Homo sapi
	1751.8	44.8	1970	6	AX431388	AX431388 Sequence
7	1711.6	43.8	161827	2	AC139668	AC139668 Papio ham
8	1448.2	37.0	179879	2	AC026883	AC026883 Homo sapi
9	778.2	19.9	190784	2	AC139669	AC139669 Lemur cat
10	500	12.8	17088	4	AX241932	AY241932 Bos tauru
11	460	11.8	164201	9	AL157935	AL157935 Human DNA
12	450.4	11.5	617	6	AX357255	AX357255 Sequence
13	447.2	11.4	200288	2	AL119725	AL119725 Homo sapi
14	433	11.1	175842	9	AL731547	AL731547 Human DNA
C 15	423.8	10.8	190508	9	AC005037	AC005037 Homo sapi
C 16	415.2	10.6	44496	9	AC004760	AC004760 Homo sapi
17	415.2	10.6	190814	9	AC006111	AC006111 Homo sapi
C 18	408.8	10.5	183690	9	AC091180	AC091180 Homo sapi
19	407.4	10.4	170528	9	AC083866	AC083866 Homo sapi
20	404.2	10.4	217615	9	AC011005	AC011005 Homo sapi
C 21	404.2	10.3	137289	9	AL161909	AL161909 Human DNA
22	403	10.3	189179	2	AC090265	AC090265 Homo sapi
C 23	402.8	10.3	181597	9	AC090515	AC090515 Homo sapi
C 24	402.8	10.3	187734	9	AC025918	AC025918 Homo sapi
C 25	401.8	10.3	204963	2	AC145617	AC145617 Homo sapi
26	394.2	10.1	179599	2	AC013278	AC013278 Homo sapi
27	394.2	10.1	195558	9	AC007448	AC007448 Homo sapi
28	392.2	10.0	155559	2	AL133336	AL133336 Homo sapi
C 29	392.2	10.0	169835	9	AP005717	AP005717 Homo sapi
C 30	391.6	10.0	171641	9	AC018512	AC018512 Homo sapi
C 31	390	10.0	139632	9	AC145334	AC145334 Pan trogl
C 32	390	10.0	171555	2	AC087537	AC087537 Homo sapi
C 33	390	10.0	181891	9	AC023356	AC023356 Homo sapi
C 34	389.8	10.0	67199	9	AC025162	AC025162 Homo sapi
C 35	388.4	9.9	172421	2	AC013279	AC013279 Homo sapi
C 36	386.2	9.9	112364	9	AC109927	AC109927 Homo sapi
C 37	386	9.9	73162	9	AC105095	AC105095 Homo sapi
C 38	386	9.9	197092	2	AC036155	AC036155 Homo sapi
C 39	384.8	9.8	140529	2	AC011486	AC011486 Homo sapi
40	384	9.8	133475	9	AC006254	AC006254 Homo sapi
41	382.4	9.8	189639	9	AC010997	AC010997 Homo sapi
C 42	381.6	9.8	80775	9	AC068539	AC068539 Homo sapi
C 43	381.2	9.7	150681	9	AC006011	AC006011 Homo sapi
44	381.2	9.7	187064	9	AC011476	AC011476 Homo sapi
45	381.2	9.7	188413	2	AC145834	AC145834 Homo sapi

ALIGNMENTS

RESULT 1	AX528729	Sequence 15 from Patent WO0236780.	4150 bp	DNA	linear	PAT 21-NOV-2002
AX528729	LOCUS	Sequence 15 from Patent WO0236780.	4150 bp	DNA	linear	PAT 21-NOV-2002
DEFINITION	AX528729	Sequence 15 from Patent WO0236780.	4150 bp	DNA	linear	PAT 21-NOV-2002
ACCESSION	AX528729	Sequence 15 from Patent WO0236780.	4150 bp	DNA	linear	PAT 21-NOV-2002
VERSION	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
KEYWORDS	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
SOURCE	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
ORGANISM	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
REFERENCE	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
AUTHORS	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
TITLE	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002
JOURNAL	AX528729.1	GI:25172800	4150 bp	DNA	linear	PAT 21-NOV-2002

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Prouty, S.M., Zhang, L. and Stenn, K.S.
Stearyl-coa desaturase gene promoter
Patent: WO 0236780-A 15 10-MAY-2002;

1981 ATACAAAAATTAGCTAAGTGTGGCGCATGCTGTAAATCCACGCTACTTGGAGGGT 2040
2041 GAGCAGAGAAATTTCTTGAACCCGGAGGAGAGGTTGCACTGAAGCGAGATCACACCA 2100
2041 GAGCAGAGAAATTTCTTGAACCCGGAGGAGAGGTTGCACTGAAGCGAGATCACACCA 2100
2101 CTGCACTCCAGCCTTGGGGGAGAGAGCGAGACTTCTCTCAAAAAAACAACCAAGA 2160
2101 CTGCACTCCAGCCTTGGGGGAGAGAGCGAGACTTCTCTCAAAAAAACAACCAAGA 2160
2161 ATTAAGCAAAATTAGACATTTGACAGAGAACTCTGAAGGGGTGACACACGACACGATTTT 2220
2161 ATTAAGCAAAATTAGACATTTGACAGAGAACTCTGAAGGGGTGACACACGACACGATTTT 2220
2221 TGTGCCATCCCAAGTACTTCTGAGGCACTCACTGGATGAGCTGTCCACATCTGAATCA 2280
2221 TGTGCCATCCCAAGTACTTCTGAGGCACTCACTGGATGAGCTGTCCACATCTGAATCA 2280
2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCCT 2340
2281 TCCAGTCTTGTTCAGAACTTTTCAACCGGACAGGGAGCCAGGACTGGAATGCACTCCT 2340
2341 GGTCACTGGCCAGAGAGTTGGCTTTGACCCCTGACAGCAGTGGCCACAAGAGGAGCTGTT 2400
2341 GGTCACTGGCCAGAGAGTTGGCTTTGACCCCTGACAGCAGTGGCCACAAGAGGAGCTGTT 2400
2401 AGTCTACTCCAGGAAATCCAGGTGCTTGTCTTCTCTGGGAAGTGAATCAATGGCGCAG 2460
2401 AGTCTACTCCAGGAAATCCAGGTGCTTGTCTTCTCTGGGAAGTGAATCAATGGCGCAG 2460
2461 CACTCCGTATTTCTCTCTCCAGGGGAAGGATCTTAGGGCAAGTATTTGGGAAAGACA 2520
2461 CACTCCGTATTTCTCTCTCCAGGGGAAGGATCTTAGGGCAAGTATTTGGGAAAGACA 2520
2521 TGGCATGGAAGACACCGGTGTAATCATAGCTGCTGCTGCTGCTGAGCTCTCATGTAA 2580
2521 TGGCATGGAAGACACCGGTGTAATCATAGCTGCTGCTGCTGCTGAGCTCTCATGTAA 2580
2581 GGCTCTACAGACACGAAAGATGGGGACAGAGGACAGATCAGTAGGGTTCAGAGCATC 2640
2581 GGCTCTACAGACACGAAAGATGGGGACAGAGGACAGATCAGTAGGGTTCAGAGCATC 2640
2641 TCAGGACCCAGGGCAATATGTTCTCAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2641 TCAGGACCCAGGGCAATATGTTCTCAGCAGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2701 GTTCTGGGCTCAACTGCCAGTCCGTCACCTTACTGCTGTGACCATGGGCAAGTTA 2760
2701 GTTCTGGGCTCAACTGCCAGTCCGTCACCTTACTGCTGTGACCATGGGCAAGTTA 2760
2761 TTCATCTCTCCATATCTCTTTCTCACTTTAAATGGAATTAATGGGGTACCCACCTCC 2820
2761 TTCATCTCTCCATATCTCTTTCTCACTTTAAATGGAATTAATGGGGTACCCACCTCC 2820
2821 CAGGGTCACAGAGGCTTACAGAAACGATTTCTGTAATTCGCTTGCAGTAATATTC 2880
2821 CAGGGTCACAGAGGCTTACAGAAACGATTTCTGTAATTCGCTTGCAGTAATATTC 2880
2881 AATACCTGCCAGCTATTTCTTATTCACATCCAGGCCCTTTGCGCTGCTGCTGGTGAATA 2940
2881 AATACCTGCCAGCTATTTCTTATTCACATCCAGGCCCTTTGCGCTGCTGCTGGTGAATA 2940
2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAAATTAACAACCTGCCAG 3000
2941 CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAAATTAACAACCTGCCAG 3000
3001 TCTGAAGATCTCCAAAACATCCGACGATCCTTGAGGCGGGGCTGGGGATGGGAC 3060
3001 TCTGAAGATCTCCAAAACATCCGACGATCCTTGAGGCGGGGCTGGGGATGGGAC 3060
3061 TGCCCGCCGGGCTCTGAAAGAGATGCTGCGGCGGACACACACACGACGACCTG 3120
3061 TGCCCGCCGGGCTCTGAAAGAGATGCTGCGGCGGACACACACACGACGACCTG 3120

3121 TGTCTGGCGCGAGTCCGCTGCGGTCCTCCGCTGAGCAGCGCTGCTGCTGGGCGGGC 3180
3121 TGTCTGGCGCGAGTCCGCTGCGGTCCTCCGCTGAGCAGCGCTGCTGCTGGGCGGGC 3180
3181 AGAGCCATTGTTCCAGGCGTACCAGCCCCCCCCCGCTGCTCCCTGCGAGGAGGCGGCT 3240
3181 AGAGCCATTGTTCCAGGCGTACCAGCCCCCCCCCGCTGCTCCCTGCGAGGAGGCGGCT 3240
3241 TCCCGCTGCTCCAGCTCCAGATCCTGGGTGGCTGCCAGCTCTCCCTGCGAGGCGCTG 3300
3241 TCCCGCTGCTCCAGCTCCAGATCCTGGGTGGCTGCCAGCTCTCCCTGCGAGGCGCTG 3300
3301 GGGGACCGGAGACGAGACGAGATGTTAGTGTGGCGGCTCCCGAGGCTTCCCACT 3360
3301 GGGGACCGGAGACGAGACGAGATGTTAGTGTGGCGGCTCCCGAGGCTTCCCACT 3360
3361 GTTTCCTGAGAACTTCCCACTGCTCCCGAGTCCCGAGTCCCGAGGCGGCT 3420
3361 GTTTCCTGAGAACTTCCCACTGCTCCCGAGTCCCGAGTCCCGAGGCGGCT 3420
3421 CCTGGCTAGGCTCCCGGCTCCAGCCCCCAACCGGCTCCCGAGGCTTCCAGAGAAA 3480
3421 CCTGGCTAGGCTCCCGGCTCCAGCCCCCAACCGGCTCCCGAGGCTTCCAGAGAAA 3480
3481 GCTCCCGACCGGGATGCTCCGCGCAGAGGCTCCAGCGGCTGGAAGAGAGTGAAGG 3540
3481 GCTCCCGACCGGGATGCTCCGCGCAGAGGCTCCAGCGGCTGGAAGAGAGTGAAGG 3540
3541 AGAAACAGAGGAGGAGGAGGAGGAGTGGCGGACAGAGGAAACAGCAGATTGCGCGA 3600
3541 AGAAACAGAGGAGGAGGAGGAGGAGTGGCGGACAGAGGAAACAGCAGATTGCGCGA 3600
3601 GCCAATGGCAACGAGGAGGAGTGGCAAAATTCCTTCCGCTGCGGCTTCCAGGCGG 3660
3601 GCCAATGGCAACGAGGAGGAGTGGCAAAATTCCTTCCGCTGCGGCTTCCAGGCGG 3660
3661 GTTTACAGAGCTCATTTAGCTATTTCCAGAGGAGGAGGAGGAGGCGGCTGG 3720
3661 GTTTACAGAGCTCATTTAGCTATTTCCAGAGGAGGAGGAGGAGGCGGCTGG 3720
3721 TGTGGTGTGGTGTGGAGCATTCCTGGCGGCTGCTGGCTGCGGCTGCGGCTGCGGCT 3780
3721 TGTGGTGTGGTGTGGAGCATTCCTGGCGGCTGCTGGCTGCGGCTGCGGCTGCGGCT 3780
3781 TCTGTCTCTCCCTCCGCTTACCTTCCAGCGGAGCCCGGCTGCGGCTGCGGCTGCGGCT 3840
3781 TCTGTCTCTCCCTCCGCTTACCTTCCAGCGGAGCCCGGCTGCGGCTGCGGCTGCGGCT 3840
3841 TCGCACTTTGCGCTTGGCAGCGATAAAGAGGAGGCTGAGAAATACCGACACGCT 3900
3841 TCGCACTTTGCGCTTGGCAGCGATAAAGAGGAGGCTGAGAAATACCGACACGCT 3900
3901 CACCCGTTGC 3910
3901 CACCCGTTGC 3910

RESULT 2
AF320307

LOCUS

DEFINITION

AF320307

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF320307 Homo sapiens stearyl-CoA desaturase (SCD) gene, exon 1 and intron 1 and partial cds.

AF320307.1 GI:14150490

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4150)

Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S.M.
Isolation and characterization of the human stearyl-CoA desaturase

1441 TTTTCTTTTCTTTTGTGAGAAAGTCTTGCTCTTTTCACCCCTGGCTGGAGGGCAGTGGTG 1500
1441 TTTTCTTTTCTTTTGTGAGAAAGTCTTGCTCTTTTCACCCCTGGCTGGAGGGCAGTGGTG 1500
1501 TGATCTCAGCTCAGTCAACTTTTCCCTCCCGGGTTCAAGCAATCTCTCTGCTCAGCCT 1560
1501 TGATCTCAGCTCAGTCAACTTTTCCCTCCCGGGTTCAAGCAATCTCTCTGCTCAGCCT 1560
1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCCCAGCTGATTTTGTATTTTAG 1620
1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCCCAGCTGATTTTGTATTTTAG 1620
1621 TAGAGATGGGATTTACATTTTGTGCGCAAGCTGGTCTCAAACTTTTGTGCTGTCATAATTG 1680
1621 TAGAGATGGGATTTACATTTTGTGCGCAAGCTGGTCTCAAACTTTTGTGCTGTCATAATTG 1680
1681 TTGTAACATATTGTTCTTTTGTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG 1740
1681 TTGTAACATATTGTTCTTTTGTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAG 1740
1741 AATCCAAATCAGTGTGTTGGTTGACCACTGTCTACCTTGAGAACCAACAGTGTGACAGGGC 1800
1741 AATCCAAATCAGTGTGTTGGTTGACCACTGTCTACCTTGAGAACCAACAGTGTGACAGGGC 1800
1801 CTCAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCCGGGCC 1860
1801 CTCAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATGAAATATTTCTCCGGGCC 1860
1861 AGGCGTGGTGGCTCATGCTGTAATCCAGCACCTTTGGGAGGCCAAGGCATGTGATCAC 1920
1861 AGGCGTGGTGGCTCATGCTGTAATCCAGCACCTTTGGGAGGCCAAGGCATGTGATCAC 1920
1921 CTGAGGTCAGAGGTTCAAAACAGCGCTGGCCAAACATGCTGAAACCCCGTCTCTATAAAA 1980
1921 CTGAGGTCAGAGGTTCAAAACAGCGCTGGCCAAACATGCTGAAACCCCGTCTCTATAAAA 1980
1981 ATACAAAATAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGT 2040
1981 ATACAAAATAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGT 2040
2041 GAGCAGAGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCAGTGAAGCGAGATCACACCA 2100
2041 GAGCAGAGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCAGTGAAGCGAGATCACACCA 2100
2101 CTGCACCTCCAGCTGGGGGAGAGCGAGACTTCCCTCTCAAAAAACAAAAACAAAAAGA 2160
2101 CTGCACCTCCAGCTGGGGGAGAGCGAGACTTCCCTCTCAAAAAACAAAAACAAAAAGA 2160
2161 ATTAAGCAAAATTAGACATTGCAGAGAGAACTGAAGGGGGTTCAGACCACTGACAGATTTC 2220
2161 ATTAAGCAAAATTAGACATTGCAGAGAGAACTGAAGGGGGTTCAGACCACTGACAGATTTC 2220
2221 TGTGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGAGCTGTCCACATCTGAAATCA 2280
2221 TGTGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGAGCTGTCCACATCTGAAATCA 2280
2281 TCCAGTCTTCTTTCAGAACTTTTCACCGGACAGGGACCGAGGCTGGAATGCGATCTCCT 2340
2281 TCCAGTCTTCTTTCAGAACTTTTCACCGGACAGGGACCGAGGCTGGAATGCGATCTCCT 2340
2341 GGTCACTGGCCAGAGAGTTGGCCCTTGACCCCTGAGACCACTGGGCCAAACAAAGAGGCTGCTT 2400
2341 GGTCACTGGCCAGAGAGTTGGCCCTTGACCCCTGAGACCACTGGGCCAAACAAAGAGGCTGCTT 2400
2401 AGTCTACCTCCAGGAAATCCAGAGTGTCTTCTCTCTGGAAAGTGAATCATTTGGCGCAG 2460
2401 AGTCTACCTCCAGGAAATCCAGAGTGTCTTCTCTCTGGAAAGTGAATCATTTGGCGCAG 2460
2461 CACTCCGTAATTTTCTCTCTCTCCAGGGAGAGTCTTAGGGCAGTATTTGGGAAAGACA 2520
2461 CACTCCGTAATTTTCTCTCTCTCCAGGGAGAGTCTTAGGGCAGTATTTGGGAAAGACA 2520
2521 TGGGCATGGAAGACACCGGGTGAATGTCATAGCTGCTGGTTCTGAGCTCTCATGGTAA 2580

2521 TGGGCATGGAAGACACCGGGTGAATGTCATAGCTGCTGGTTCTGAGCTCTCATGGTAA 2580
2581 GGCTCTCTACACACCGGAAAGATGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC 2640
2581 GGCTCTCTACACACCGGAAAGATGGGGCACAGGACAGATCAGTAGGGTCAGAGCATC 2640
2641 TCAGGGACCGAGGGCAATATGTTCTGTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2641 TCAGGGACCGAGGGCAATATGTTCTGTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGT 2700
2701 GTTTCTGGGCTCAACTGCCAGCTCGTCTACCTTACTGTTGCTGTGACCATGGGCAAGTTA 2760
2701 GTTTCTGGGCTCAACTGCCAGCTCGTCTACCTTACTTCTTCTTAAATGGAATTAAGAGCTTGGGCTTAA 2760
2761 TTCCATCTCTCATATCTCTTCTCTCACTTTTAAATGGAATTAAGAGCTTGGGCTCTCATATGGT 2820
2761 TTCCATCTCTCATATCTCTTCTCTCACTTTTAAATGGAATTAAGAGCTTGGGCTTAA 2820
2821 CAGGTCACAGAGGGCTTACAGAAAACGATTTCTGTGAATTTGGCTTGCAGTAAATTAATTC 2880
2821 CAGGTCACAGAGGGCTTACAGAAAACGATTTCTGTGAATTTGGCTTGCAGTAAATTAATTC 2880
2881 AATACCTGCCAGCTATTTCTTATTTCCACATCCAAAGCCCTTTCCGCTGCTGTGGGTGAAA 2940
2881 AATACCTGCCAGCTATTTCTTATTTCCACATCCAAAGCCCTTTCCGCTGCTGTGGGTGAAA 2940
2941 CACATGTCAGTGTCTCTGAGCGGTTTCCAAAGAGATTTCCAAATTTACAACTTGCAG 3000
2941 CACATGTCAGTGTCTCTGAGCGGTTTCCAAAGAGATTTCCAAATTTACAACTTGCAG 3000
3001 TCTGAGAAATCTCAAAAACATTCGCGACGATTCGCGCAGGCGGCTTGGGGATGGGAC 3060
3001 TCTGAGAAATCTCAAAAACATTCGCGACGATTCGCGCAGGCGGCTTGGGGATGGGAC 3060
3061 TGCCCGCCCGGGTCTCTCAACAGGATCGTGGCGCAGGACACACACAGCCAGCGCTG 3120
3061 TGCCCGCCCGGGTCTCTCAACAGGATCGTGGCGCAGGACACACACAGCCAGCGCTG 3120
3121 TGTGTGCGCGCGGAGTCCGGTCCCGGGTTCAGCAGCGCTGGCTGTGGCGCGGGC 3180
3121 TGTGTGCGCGCGGAGTCCGGTCCCGGGTTCAGCAGCGCTGGCTGTGGCGCGGGC 3180
3181 AGAGCCATTTGTTCCGAGCGCTTACGAGCCCCCGCTGCGCCGGGAGGAGCGGGCT 3240
3181 AGAGCCATTTGTTCCGAGCGCTTACGAGCCCCCGCTGCGCCGGGAGGAGCGGGCT 3240
3241 TCCCGGCTCCCAAGCTCCAGATCTCTGGGGTGGCTGCGAGCTCTCCCTGCGACCGCGCTG 3300
3241 TCCCGGCTCCCAAGCTCCAGATCTCTGGGGTGGCTGCGAGCTCTCCCTGCGACCGCGCTG 3300
3301 GGGGGAACCGGGAACCGGACCGGAGATGTTAGTGTGGGCGCCCGCCGAGGGTTCCACCACT 3360
3301 GGGGGAACCGGGAACCGGACCGGAGATGTTAGTGTGGGCGCCCGCCGAGGGTTCCACCACT 3360
3361 GTTTCTCTGAGAAATTTTCCCAAGTTCGCCACCCCTTCTCCGCTGTCGCCGAGGGCGGT 3420
3361 GTTTCTCTGAGAAATTTTCCCAAGTTCGCCACCCCTTCTCCGCTGTCGCCGAGGGCGGT 3420
3421 CTTGGGCTAGGCTTCCCGGCCCGCCAGCCCCAACCCGGGTCCCGCTCCCGCTTCCAGAGGAAA 3480
3421 CTTGGGCTAGGCTTCCCGGCCCGCCAGCCCCAACCCGGGTCCCGCTCCCGCTTCCAGAGGAAA 3480
3481 GCTTCCGACGCGGGATTCGCGGGCAGAGGCCCGCGGGTGGAAAGAGCTGAGAAGG 3540
3481 GCTTCCGACGCGGGATTCGCGGGCAGAGGCCCGCGGGTGGAAAGAGCTGAGAAGG 3540
3541 AGAAACAGAGGGGAGGGGAGCGAGAGCTGCGCGCAGAGGGAACAGCAGATTGCCCGCA 3600
3541 AGAAACAGAGGGGAGGGGAGCGAGAGCTGCGCGCAGAGGGAACAGCAGATTGCCCGCA 3600
3601 GCCAATGGCAACCGGACGAGAGGTTGGCACCAATTTCCCTTCCGCCCAATGACGAGCCGGA 3660

```

Db      3601  GCCAATGGCAACGCGCAGCAGGTGGCGACCAAAATTCCTTCGGCCCAATGACGAGCGCGA 3660
QY      3661  GTTTACAGAGCCTCATTTAGCATTTCCCGAGAGCGAGGGCAGAGGCGCGGTGG 3720
Db      3661  GTTTACAGAGCCTCATTTAGCATTTCCCGAGAGCGAGGGCAGAGGCGCGGTGG 3720
QY      3721  TGTGGTGTGGTGTGGCAGCATCCCCGGGGCCCTGCTGGGGTGGCGCGGAGCCTCGGCC 3780
Db      3721  TGTGGTGTGGTGTGGCAGCATCCCCGGGGCCCTGCTGGGGTGGCGCGGAGCCTCGGCC 3780
QY      3781  TCTGTCTCTCCCTCCCGCCCTTACTCGACGCGGAGCGCCCGCGGCGAGCACTCC 3840
Db      3781  TCTGTCTCTCCCTCCCGCCCTTACTCGACGCGGAGCGCCCGCGGCGAGCACTCC 3840
QY      3841  TCGCACTTTGCCCTGCTTCGCGAGCGGATAAAGGGGGCTGAGGAATACCGGACGCT 3900
Db      3841  TCGCACTTTGCCCTGCTTCGCGAGCGGATAAAGGGGGCTGAGGAATACCGGACGCT 3900
QY      3901  CACCCGTTGC 3910
Db      3901  CACCCGTTGC 3910

RESULT 3
AC018783/c
LOCUS      159351 bp      DNA      linear      HTG 14-MAR-2000
DEFINITION Homo sapiens clone RP11-2M13, WORKING DRAFT SEQUENCE, 12 unordered
            pieces.
ACCESSION  AC018783
VERSION     AC018783.3 GI:7223991
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 159351)
AUTHORS   Birren,B., Linton,L., Nussbaum,C. and Lander,E.
TITLE     Homo sapiens, clone RP11-2M13
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 159351)
AUTHORS   Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
            O'Connor,T., O'Donnell,P., Olivier,T.M., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
            Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
            and Zody,M.
TITLE     Direct Submission
JOURNAL   Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Mar 12, 2000 this sequence version replaced gi:6692332.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L2667
            Center clone name: 2_M13
            ----- Summary Statistics

```

```

Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151339 bases at least Q40
Consensus quality: 155177 bases at least Q30
Consensus quality: 157011 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 158251; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      2297: contig of 2297 bp in length
*      2298      2397: gap of 100 bp
*      2398      5719: contig of 3322 bp in length
*      5720      5819: gap of 100 bp
*      5820      10149: contig of 4330 bp in length
*      10150      10249: gap of 100 bp
*      10250      16496: contig of 6247 bp in length
*      16497      16596: gap of 100 bp
*      16597      21823: contig of 5227 bp in length
*      21824      21923: gap of 100 bp
*      21924      25747: contig of 3824 bp in length
*      25748      25847: gap of 100 bp
*      25848      31971: contig of 6124 bp in length
*      31972      32071: gap of 100 bp
*      32072      44449: contig of 12378 bp in length
*      44450      44549: gap of 100 bp
*      44550      64978: contig of 20429 bp in length
*      64979      65079: gap of 100 bp
*      65079      91491: contig of 26413 bp in length
*      91492      121937: contig of 30346 bp in length
*      121938      122038: gap of 100 bp
*      122038      159351: contig of 37314 bp in length.
*
* Location/Qualifiers
* 1..159351
*   /organism="Homo sapiens"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:9606"
*   /clone="RP11-2M13"
*   /clone_lib="RPC1-11 Human Male BAC"
*   1..2297
*     /note="assembly_fragment"
*     2398..5719
*       /note="assembly_fragment"
*       5820..10149
*         /note="assembly_fragment"
*         10250..16496
*           /note="assembly_fragment"
*           16597..21823
*             /note="assembly_fragment"
*             21924..25747
*               /note="assembly_fragment"
*               clone_end:SP6
*               vector_side:right
*               25848..31971
*                 /note="assembly_fragment"
*                 clone_end:17
*                 vector_side:left
*                 32072..44449
*                   /note="assembly_fragment"
*                   44550..64978
*                     /note="assembly_fragment"
*                     65079..91491
*                       /note="assembly_fragment"

```


[illegible]

QY	2043	GGCAGGAGAAATTTCTTGAACCCGGGAGCGAGAGGTTGCAGTGAAGCGACATCACACCACT	2107
Db	57350	GGCAGGAGAAATTTCTTGAACCCGGGAGCGAGAGGTTGCAGTGAAGCGACATCACACCACT	57291
QY	2103	GCACCTCAGCCTTGGGGAGAGACGGAGACTTCCTCTCAAAAAACAACAAAAA	2162
Db	57290	GCACCTCAGCCTTGGGGAGAGACGGAGACTTCCTCTCAAAAAACAACAAAAA	57231
QY	2163	TAAGCAAAATTAGACATTTGCAGAGAGAACCTTGAGGGGGTGCAGACACGTCGATTTCTG	2222
Db	57230	TAAGCAAAATTAGACATTTGCAGAGAGAACCTTGAGGGGGTGCAGACACGTCGATTTCTG	57171
QY	2223	TGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGGATGAGCTCTCAACATCTGAATCATC	2282
Db	57170	TGCCACATGCCAAGTACTTCTGAGGCATGACTGGATGGATGAGCTCTCAACATCTGAATCATC	57111
QY	2283	CAGTCTTGTTTCAGAACTTTTCACACCGGACAGGAGCGAGGACTGGAATGCGAGTCTCTGG	2342
Db	57110	CAGTCTTGTTTCAGAACTTTTCACACCGGACAGGAGCGAGGACTGGAATGCGAGTCTCTGG	57051
QY	2343	TCACCTGCCACAGAGTTGGCTTGACCCCTGAGACCAAGTGGCCCAACAAAGGAGCTGCTTAG	2402
Db	57050	TCACCTGCCACAGAGTTGGCC-TGACCCCTGAGACCAAGTGGCCCAACAAAGGAGCTGCTTAG	56992
QY	2403	TCTACCTCCACAGAAATCCAGAGTGCTGTCTTCTCTGGGAAATGAATCATTTGGCGCGACGA	2462
Db	56991	TCTACCTCCACAGAAATCCAGAGTGCTGTCTTCTCTGGGAAATGAATCATTTGGCGCGACGA	56932
QY	2463	CTCCGTAATTTCTCTCTTTCCACAGGGGAAGATCTTAGGGCAGTATTTGGGAAAGACATG	2522
Db	56931	CTCCGTAATTTCTCTCTTTCCACAGGGGAAGATCTTAGGGCAGTATTTGGGAAAGACATG	56872
QY	2523	GGCATGGAAGGACACCGGGTGAATGCATAGCCTGCTGCTGTCTGAGCTCTCATGGTAAAG	2582
Db	56871	GGCATGGAAGGACACCGGGTGAATGCATAGCCTGCTGCTGTCTGAGCTCTCATGGTAAAG	56812
QY	2583	CTTCTCAGACACGGAAAGATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATCTC	2642
Db	56811	CTTCTCAGACACGGAAAGATGGGGGCACAGGACAGATCAGTAGGGTCAGAGCATCTC	56752
QY	2643	AGGGACCGAGGGCAATATGGTCTTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGTGT	2702
Db	56751	AGGGACCGAGGGCAATATGGTCTTGAGCAGGGATTAAGAGCTTGGGCTCTCATATGGTGT	56692
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGTCATTACTACTGGTTGCTGTGACCATGGGCAAGTATT	2762
Db	56691	TTCTGGGCTCAACTGCCAGCTCCATCACTTACTACTGGTTGCTGTGACCATGGGCAAGTATT	56632
QY	2763	CCATCTCTCATATCTCTTTCTCCTCACTTTTAAATGGAATAATGGGGTACCCACCTCCCA	2822
Db	56631	CCATCTCTCATATCTCTTTCTCCTCACTTTTAAATGGAATAATGGGGTACCCACCTCCCA	56572
QY	2823	GGGTACAGAGAGGCTTACAGAAACGATTTGTGAAATTTGGCTTGGCAGTAATAATCAA	2882
Db	56571	GGGTACAGAGAGGCTTACAGAAACGATTTGTGAAATTTGGCTTGGCAGTAATAATCAA	56512
QY	2883	TACCTGCCAGCTATTCTTATTTCCACATCCAAAGCCCTTTCCCTCTGCTCTGGGGTGAACA	2942
Db	56511	TACCTGCCAGCTATTCTTATTTCCACATCCAAAGCCCTTTCCCTCTGCTCTGGGGTGAACA	56452
QY	2943	CATGTCAGTGTCTTCTGACCGTTTTCCAACAAAGAGATTTCCAAATTTACAACCTGCCAGTC	3002
Db	56451	CATGTCAGTGTCTTCTGACCGTTTTCCAACAAAGAGATTTCCAAATTTACAACCTGCCAGTC	56392
QY	3003	TGAAGATCTCCAAAACATCCCGCACGCTCTTGAGGCGCGGGCTTGGGATGGGACTG	3062
Db	56391	TGAAGATCTCCAAAACATCCCGCACGCTCTTGAGGCGCGGGCTTGGGATGGGACTG	56332
QY	3063	CCCGCCCGGGTCTTGAACAGGATGCGTGGCGGACAGCACACACACGACGACGCTGTG	3122
Db	56331	CCCGCCCGGGTCTTGAACAGGATGCGTGGCGGACAGCACACACACGACGACGCTGTG	56272
QY	3123	TGTGCGCGGAGTCTCGGTGCGGTCCGGGTGAGACGCGCTGGCTGTGGCGCGGGCGAG	3182

Db	56271	TGTGCGCGGAGTCCGGTGCCTCCGGGTGACAGCGCGTGGCTGGTGGCGGGCAG	56212
QY	3183	AGCCATTGTCGACGGCGTACCGAGCCCCCGCGCTCGCCGGAGGAGCGGGCTTC	3242
Db	56211	AGCCATTGTCGACGGCGTACCGAGCCCCCGCGCTCGCCGGAGGAGCGGGCTTC	56152
QY	3243	CCGCGT-CCCAAAGCTCCAGATCTCGGGTGGCTGCCAGTCTCCCTGCCACGCGCTGG	3301
Db	56151	CCGCGTCCCCCAAGTCCAGATCTCGGGTGGCTGCCAGTCTCCCTGCCACGCGCTGG	56092
QY	3302	GGGGA CGGGAAGACGGGACCGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTCACCACTG	3361
Db	56091	GGGGA CGGGAAGACGGGACCGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTCACCACTG	56032
QY	3362	TTTCTCTGAGAACTTCCCAAGTGCCCAACCCAGTCTCGTGTGCGGAGGCGCGTC	3421
Db	56031	TTTCTCTGAGAACTTCCCAAGTGCCCAACCCAGTCTCGTGTGCGGAGGCGCGTC	55972
QY	3422	CTGGGCTAGGCTCCGCGCCCCAGCCCCAAACCGGGTCCCAGACCCCTTCCAGAGGAAG	3481
Db	55971	CTGGGCTAGGCTCCGCGCCCCAGCCCCAAACCGGGTCCCAGACCCCTTCCAGAGGAAG	55912
QY	3482	CTCCCAACGCGGGATCCGCGCACAGGCCAGCGCGGGTGGGAAGAAGCTGAGAAGGA	3541
Db	55911	CTCCCAACGCGGGATCCGCGCACAGGCCAGCGCGGGTGGGAAGAAGCTGAGAAGGA	55852
QY	3542	GAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGCAGAGGGAACAGCAGATTGCGCGAG	3601
Db	55851	GAAACAGAGGGGAGGGGAGCGAGAGCTGGCGGCAGAGGGAACAGCAGATTGCGCGAG	55792
QY	3602	CCAAATGGCAACGCGACGAGGTGCACCAATTCCTTCCGCCAATGACAGACCGGAG	3661
Db	55791	CCAAATGGCAACGCGACGAGGTGCACCAATTCCTTCCGCCAATGACAGACCGGAG	55732
QY	3662	TTTACAGAAAGCTCATTTAGCATTTCCCAAGAGCAGGGCAGAGGCGCGGTGGT	3721
Db	55731	TTTACAGAAAGCTCATTTAGCATTTCCCAAGAGCAGGGCAGAGGCGCGGTGGT	55672
QY	3722	GTGGTGTGCGTGTGGGACGATCCCGGGCGCCTGTGCGGTGCGCGCAGAGCTCGGCCT	3781
Db	55671	GTGGTGTGCGTGTGGGACGATCCCGGGCGCCTGTGCGGTGCGCGCAGAGCTCGGCCT	55612
QY	3782	CTGTCTCTCCCTCCCGCCCTTACCTTCACGCGGGAACGCGCGGCGCAGTCACTCT	3841
Db	55611	CTGTCTCTCCCTCCCGCCCTTACCTTCACGCGGGAACGCGCGGCGCAGTCACTCT	55552
QY	3842	CGCACTTTGCCCTGCTTGGCAGCGGATAAAGGGGGCTGAGGAATAACCGACACGGTC	3901
Db	55551	CGCACTTTGCCCTGCTTGGCAGCGGATAAAGGGGGCTGAGGAATAACCGACACGGTC	55492
QY	3902	ACCGTTGCG	3910
Db	55491	ACCGTTGCG	55483

RESULT 4	
AL1139819	
LOCUS	115278 bp DNA linear PRI 18-SEP-2001
DEFINITION	Human DNA sequence from clone RP11-34D15 on chromosome 10, complete sequence.
ACCESSION	AL1139819
VERSION	AL1139819.8 GI:15705162
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 115278)
AUTHORS	Tracey, A.
TITLE	Direct Submission
JOURNAL	Submitted (18-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerv@sanger.ac.uk Clone

requests: clonerquest@sanger.ac.uk

On Sep 19, 2001 this sequence version replaced gi:14586033.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr10>

RP11-34D15 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-34D15 It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-34D15 is at 115278 in this sequence. The true left end of clone RP11-285F16 is at 56886 in this sequence. The true right end of clone RP11-316M21 is at 200 in this sequence.

ORIGIN

Query Match	99.0%;	Score 3869.8;	DB 9;	Length 115278;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3904;	Conservative 0;	Mismatches 2;	Indels 3;	Gaps 3;
QY	3	GAATTCATCCATTTAAATCATACAAATTTAAATGGCTTTTAGTATATATTCACAGGTGTGCAT	62	
Db	64592	GAATTCATCCATTTAAATCATACAAATTTAAATGGCTTTTAGTATATATTCACAGGTGTGCAT	64651	
QY	63	CCATCAAAATCCATTTTAGAACAGTTTTATTACTCCAAAATAAAACCCCTGCATTCCTTAG	122	
Db	64652	CCATCAAAATCCATTTTAGAACAGTTTTATTACTCCAAAATAAAACCCCTGCATTCCTTAG	64711	
QY	123	CCATGACCCGCCAAACATCTCCATTCCTTCGAGCCCTGGGCAACCAACCAATCTACTT	182	
Db	64712	CCATGACCCGCCAAACATCTCCATTCCTTCGAGCCCTGGGCAACCAACCAATCTACTT	64771	
QY	183	TCGTCTCTATAAAATTGGCAATTCGGAATTTGGAATTTTCATATAAATGAAGCAACCAACATGT	242	
Db	64772	TCGTCTCTATAAAATTGGCAATTCGGAATTTGGAATTTTCATATAAATGAAGCAACCAACATGT	64831	
QY	243	GAGACTTTGTGACTGGCTGCTTTCACTTAGCATTCCTATTTTAAAGGCTCATTTATGTTACA	302	
Db	64832	GAGACTTTGTGACTGGCTGCTTTCACTTAGCATTCCTATTTTAAAGGCTCATTTATGTTACA	64891	
QY	303	GTACTTAGCAGTACTTCATTCCTTTTTTATCTCAAATGGTATTCACATGTGTGGGTATCC	362	
Db	64892	GTACTTAGCAGTACTTCATTCCTTTTTTATCTCAAATGGTATTCACATGTGTGGGTATCC	64951	


```

QY 3662 TTTACAGAGCCTCATTAGCATTTCCACAGAGGAGGGGAGGGGAGAGGCGGGTGGT 3721
Db 68250 TTTACAGAGCCTCATTAGCATTTCCACAGAGGAGGGGAGGGGAGAGGCGGGTGGT 68309
QY 3722 GTGGTGTGGTGTGGGAGCATCCCGGCGCCCTGTCTGCGGTGCGCGGAGCCTCGGCGCT 3781
Db 68310 GTGGTGTGGTGTGGGAGCATCCCGGCGCCCTGTCTGCGGTGCGCGGAGCCTCGGCGCT 68369
QY 3782 CTGTCTCTCTCCCTCCCGCCCTTACTCTCCACGCGGAGCGCGCGCGGAGTCACTCTCT 3841
Db 68370 CTGTCTCTCTCCCTCCCGCCCTTACTCTCCACGCGGAGCGCGCGCGGAGTCACTCTCT 68429
QY 3842 CGGACTTTGCGCCTGTGTGCGACGGATAAAGGGGGCTGAGGAATACCGGACACGGTC 3901
Db 68430 CGGACTTTGCGCCTGTGTGCGACGGATAAAGGGGGCTGAGGAATACCGGACACGGTC 68489
QY 3902 ACCCGTTGC 3910
Db 68490 ACCCGTTGC 68498

RESULT 5
AL391723/c
LOCUS AL391723 185644 bp DNA linear HTG 11-SEP-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-38B21, WORKING DRAFT
ACCESSION AL391723
VERSION AL391723.7 GI:10186824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burton, J.
Direct Submission
Submitted (10-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Sep 19, 2000 this sequence version replaced gi:10178830.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bA38B21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 175002 bases at least Q40
Consensus quality: 179431 bases at least Q30
Consensus quality: 182156 bases at least Q20
Insert size: 184344; sum-of-contigs
Insert size: 140788; 6.0% error; agarose-fp
Quality coverage: 4.15x in Q20 bases; sum-of-contigs Quality
coverage: 6.51x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 59110: contig of 59110 bp in length
* 59111 59210: gap of 100 bp
* 59211 61285: contig of 2075 bp in length
* 61286 61386 64379: gap of 100 bp
* 61386 64379: contig of 2994 bp in length

```

```

* 64380 64479: gap of 100 bp
* 64480 66747: contig of 2268 bp in length
* 66747 66848 66847: gap of 100 bp
* 66848 114355: contig of 47508 bp in length
* 114355 114456 114455: gap of 100 bp
* 114456 125406: contig of 10951 bp in length
* 125406 125507 125506: gap of 100 bp
* 125507 127673: contig of 2167 bp in length
* 127673 127774 127773: gap of 100 bp
* 127774 130111 130110: contig of 2338 bp in length
* 130111 130212 130211: gap of 100 bp
* 130212 132354: contig of 2143 bp in length
* 132354 132456 132454: gap of 100 bp
* 132456 136537 136536: contig of 4083 bp in length
* 136537 136638 136637: gap of 100 bp
* 136638 139127 139126: contig of 2490 bp in length
* 139127 139228 139227: gap of 100 bp
* 139228 181356 181355: contig of 42128 bp in length
* 181356 181456 181455: gap of 100 bp
* 181456 183525 183524: contig of 2069 bp in length
* 183525 185644 185643: gap of 100 bp
* 185644 185644: contig of 2020 bp in length.

FEATURES
            Location/Qualifiers
     source          1..185644
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-38B21"
                     /clone_lib="RPC1-11.1"
     misc_feature    1..59110
                     /note="assembly_fragment:02307"
                     clone_end:17
                     vector_side:left
                     59211..61285
                     /note="assembly_fragment:00197"
                     fragment_chain:1
                     61386..64379
                     /note="assembly_fragment:01916"
                     fragment_chain:1
                     64480..66747
                     /note="assembly_fragment:00079"
                     fragment_chain:1
                     66848..114355
                     /note="assembly_fragment:00585"
                     fragment_chain:2
                     114456..125406
                     /note="assembly_fragment:02383"
                     fragment_chain:2
                     125507..127673
                     /note="assembly_fragment:00150"
                     127774..130111
                     /note="assembly_fragment:00464"
                     130212..132354
                     /note="assembly_fragment:00629"
                     132455..136537
                     /note="assembly_fragment:00814"
                     136638..139127
                     /note="assembly_fragment:00861"
                     139228..181355
                     /note="assembly_fragment:02212"
                     181456..183524
                     /note="assembly_fragment:02327"
                     183625..185644
                     /note="assembly_fragment:02419"

ORIGIN
Query Match          45.5%; Score 1779.2; DB 2; Length 185644;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1823; Conservative 0; Mismatches 118; Indels 3; Gaps 3;
QY 1968 GTCTCTACTAAATAACAAAATAATTAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGC 2027
|||||
|||||
|||||
|||||

```

[illegible][illegible]

RESULT 6					
AX431388					
LOCUS					
DEFINITION					
AX431388					
ACCESSION					
AX431388.1					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					

```

same
JOURNAL      Patent: WO 0240666-A 21 23-MAY-2002,2
XENON GENETICS INC (CA)
FEATURES     Location/qualifiers
              1. 1970
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
source

```

ORIGIN

Query Match	44.8%;	Score 1751.8;	DB 6;	Length 1970;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1786;	Conservative 0;	Mismatches 2;	Indels 3;	Gaps 3;
QY	2121	GAGAGCGAGACTTCCTCTCAAAAACAAACAAACAAAGAAATTAAGCAAATTTAGACATTG	2180	
DB	1	GAGAGCGAGACTTCCTCTCAAAAACAAACAAACAAAGAAATTTAGCAAATTTAGACATTG	60	
QY	2181	CAGAGAAACCTGAAAGGGGTGAGACCACTGACAGATTTCTGTGCCACATGCCAAGTACT	2240	
DB	61	CAGAGAGAACCTGAAAGGGGGTCAAGCACTGACAGATTTCTGTGCCACATGCCAAGTACT	120	
QY	2241	TCTGAGGCATGACTGGATGAGCTTCACATCTGAAATCATTCAGTCTTGTTCAGAACTT	2300	
DB	121	TCTGAGGCATGACTGGATGAGCTTCACATCTGAAATCATTCAGTCTTGTTCAGAACTT	180	
QY	2301	TCACACGGGACGGAGCCAGGACTGGAATGTCAGTCTCCTGGTCTCACTGGGCCAGAGAGTTG	2360	
DB	181	TCACACGGGACGGAGCCAGGACTGGATGTCAGTCTCCTGGTCTCACTGGGCCAGAGAGTTG	240	
QY	2361	GCCTTGACCCCTGAGACGAGTGGCCAAACAAAGGAGCTGCTTAGTCTACCTCCCAAGGAAATC	2420	
DB	241	GCC-TGACCCCTGAGACGAGTGGCCAAACAAAGGAGCTGCTTAGTCTACCTCCCAAGGAAATC	299	
QY	2421	CGAGGTGCTTGCTTCTCTGGGAAGTGAATCAATTTGGGGCGACACTCCGTATTTTCTCCTCT	2480	
DB	300	CGAGGTGCTTGCTTCTCTGGGAAGTGAATCAATTTGGGGCGACACTCCGTATTTTCTCCTCT	359	
QY	2481	TCCACGGGGAAGGATCCTTAGGGCAGATTAATTTGGGAAAGACATGGGCATGGAAGGACACCCGG	2540	
DB	360	TCCACGGGGAAGGATCCTTAGGGCAGATTAATTTGGGAAAGACATGGGCATGGAAGGACACCCGG	419	
QY	2541	GTGAATGCAATAGCCTGCCTGGTTCTGAGCTCTCATGTTAAGGTCTCTPACAGACACGGAAA	2600	
DB	420	GTGAATGCAATAGCCTGCCTGGTTCTGAGCTCTCATGTTAAGGTCTCTPACAGACACGGAAA	479	
QY	2601	AGATGGGGGCACAGGACACAGATCAGTAGGGTCAGAGCATCTCAGGAGCCAGGGCCATAT	2660	
DB	480	AGATGGGGGCACAGGACACAGATCAGTAGGGTCAGAGCATCTCAGGAGCCAGGGCCATAT	539	
QY	2661	GGTCTCTGACGAGGGATTAAGAGCTTGGGGCTCTCATATGGTGTTCCTGGGCTCAACTGCCA	2720	
DB	540	GGTCTCTGACGAGGGATTAAGAGCTTGGGGCTCTCATATGGTGTTCCTGGGCTCAACTGCCA	599	
QY	2721	GCTCCGTCACTTACTGGTTGCTGTGACCAATGGGCAAGTTATTCCCATCTCTCCATATCTCT	2780	
DB	600	GCTCCCATCTTACTGGTTGCTGTGACCAATGGGCAAGTTATTCCCATCTCTCCATATCTCT	659	
QY	2781	TTTCCTCACTTTTAAATGGAATGCGGTACCCACTCCAGGGTCAACAGAGGCTTA	2840	
DB	660	TTTCCTCACTTTTAAATGGAATGCGGTACCCACTCCAGGGTCAACAGAGGCTTA	719	
QY	2841	CAGAAAACGATTTCTGTGAATTGGCTTGCAGTAATAAATTCATACCTGCCAGCTATTCCT	2900	
DB	720	CAGAAAACGATTTCTGTGAATTGGCTTGCAGTAATAAATTCATACCTGCCAGCTATTCCT	779	
QY	2901	ATTTCGACATCCAAAGCCCTTTGGCCTGCTGTGGTGAAAACAATGTCAGTGTTCCTCGA	2960	
DB	780	ATTTCGACATCCAAAGCCCTTTGGCCTGCTGTGGTGAAAACAATGTCAGTGTTCCTCGA	839	
QY	2961	CGGTTTCCCAAGAAGATTTCCAAATTAACACCTGCCAGTCTGAAGATCTCCAAAAACA	3020	
DB	840	CGGTTTCCCAAGAAGATTTCCAAATTAACACCTGCCAGTCTGAAGATCTCCAAAAACA	899	

RESULT 7
AC139668

AC139668	161827 bp	DNA	linear	HTG 08-FEB-2003
LOCUS				
DEFINITION	Papio hamadryas clone RP41-194B17, WORKING DRAFT SEQUENCE.			
ACCESSION	AC139668			
VERSION	AC139668.1	GI:28273351		
KEYWORDS	HTG; HTGS; PHASE1; HTGS DRAFT.			
SOURCE	Papio hamadryas (hamadryas baboon)			
ORGANISM	Papio hamadryas			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
	Papio			

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Cercopithecinae; Papio.
1 (bases 1 to 161827)
Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.
Direct Submission
Unpublished
2 (bases 1 to 161827)
Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng, Z., Malinov, I. and Rubin, E.M.
Direct Submission
Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: EGAHERK
Center Project Name: B051
Bac Clone Name: RP41-194B17

This sequence has been compared to sequences of other species
using Vista (<http://www.gsdlbl.gov/VISTA>). The results can be
viewed at:
http://pga.lbl.gov/cgi-bin/search_cvgcd?type=nk&value=SCD

The order-orientation of the draft sequence was accomplished by
using:
Avid (<http://baboon.math.berkeley.edu/mavid/>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:
Sequencing vector: Plasmid; pUC18
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap version 0.990329.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 161827: contig of 161827 bp in length.
Location/Qualifiers
1. 161827
/organism="Papio hamadryas"
/mol_type="genomic DNA"
/db_xref="taxon:9557"
/clone="RP41-194B17"

FEATURES
source

Query Match 43.8%; Score 1711.6; DB 2; Length 161827;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1896; Conservative 0; Mismatches 134; Indels 24; Gaps 8;

QY 1858 GCACGGCTGGTGGCTCATCGCTGTAATCCAGCACTTTGGAGGGCCAGGCGATGGAT 1917
Db 139308 GCTGGGCACGGTGGCTCATCGCTGTAATCCAGCACTTTGGAGGGCTAAGAGGGCGAT 139367

QY 1918 CACCTGAGGTCAGAGGTTCAAAACCGAGCTGGCCAAATGGTGAACCCCGCTCTACTA 1977
Db 139368 CA--TGAGGTCAGGAATGCAAGACCGAGCTGACCAATAGGGTGAACCCCGCTGACTA 139425

QY 1978 AAAATCAAAAAATTAGCTAAGTGTTGGTGGCGATGCTGTAATCCAGCTACTTGGGAG 2037
Db 139426 AAAATCAAAAAATTAGCTAAGTGTTGGTGGCGATGCTGTAATCCAGCTACTTGGGAG 139485

QY 2038 GGTGAGCGAGGAAATTTCTTGAACCGGAGGAGGAGGTTGAGTGAAGCGAGATCA 2097
Db 139486 GCTGAGCGAGGAAATTTCTTGAACCGGAGGAGGTTGAGTGAAGCGAGATCA 139545

QY 2098 CCACCTGCATCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2157

Db 139546 CCACCTGCATCCAGCTGGGGACAGAGGACTCCCTCTCAAAAAACAAAAACAA 139605
QY 2158 AGAATTAAGCAAAATTAGACATTCGAGAGAGACCTGAAGGGGTGAGACAGTACAGAT 2217
Db 139606 ATAAATCAGAAATTTAGACATTCGAGAGAGAACCTGAAGGGGAGAGACAGTACAGAT 139665
QY 2218 TTCTGTGCCATGCAATGCAATGCTTCTGAGGATGATGATGAGCTGTCCACATCTGAAA 2277
Db 139666 TTCTGTGCCATGCAATGCAATGCTTCTGAGGATGATGATGAGCTGTCCACATCTGAAA 139725
QY 2278 TCATCCAGTCTGTTTCAGAACTTTTCACCCGAGAGGAGCCAGGACTGGAATGAGTCT 2337
Db 139726 TCATCCAGTCTGTTTCAGAACTTTTCACCCGAGAGGAGCCAGGACTGGAATGAGTCT 139785
QY 2338 CCTGTCTACTGGCCAGAGATTGGCTTACCTCTGAGACCACTGAGTGGCCAAAGAGAGTG 2397
Db 139786 CCTGTCTACTGGCCAGAGATTGGCTTACCTCTGAGACCACTGAGTGGCCAAAGAGAGTG 139844
QY 2398 CTTAGTCTACTCCCGAGGAAATCCAGGTGCTTCTCTCTGGGAAGTGAATCAATTGCGG 2457
Db 139845 ATTAGTCTTCTCTCCAGGAAATCCAGG-----TGGGAGCGAATCAATTAGCG 139892
QY 2458 CAGCACTCCGTATTTCTCTCTTCCAGGGGAAGATCTTGGGCAAGTATTGGGAAAG 2517
Db 139893 CAGCACTCCGTATTTCTCTCTTCCAGGGAATGATCTTGGGCAAGTATTGGGCAAG 139952
QY 2518 ACATGGGCTGGAAGAGACACCGGGTGAATGATGATGAGTGGTCTGAGCTCTCATGG 2577
Db 139953 ACATGGGCTGGAAGAGAGAGGGTGAATGATGATGAGTGGTCTGAGCTCTCATGG 140012
QY 2578 TAAGGCTCTACAGACACGGAAGATGAGGGGACACAGGACAGATCAGTAGGGTCAAGC 2637
Db 140013 TAAGACTCTCTACAGACACCGGAAAGAGGGGGGACACAGGACAGATCAGTAGGGTCAAGC 140072
QY 2638 ATCTCGGGACCGGAGCAATATGCTCTCAGCAGGGAATTAAGAGCTTGGGCTCTCATAT 2697
Db 140073 ATCTCGGGACCGGAGCAATATGCTCTCAGCAGGGAATTAAGAGCTTGGGCTCTCATAT 140132
QY 2698 GGTGTTTCTGGGCTCAACTGCCAGCTCCGTCATCTACTGCTGCTGACATGGGCAAG 2757
Db 140133 GGCATTTCTGGGCTCAATTGCAGCTCCATCACTTACTGCTGCTGATCATGGGCAAG 140192
QY 2758 TTATTCATCTCTCATATCTCTTCTCTCACTTTTAAATGGAATTAATGGGTACCCACC 2817
Db 140193 TTATTCATCTCTCATATCTCTTCTCTCACTTTTAAATGGAATTAATGGGTACCCACC 140252
QY 2818 TCCACGGCTCAGAGAGGCTTACAGAAACGATTTCTTGAATTTGGCTTGCAGTAAATA 2877
Db 140253 TCCACGGCTCAGAGAGGATTAAGAAACGATTTCTTGAATTTGGCTTGCAGTAAATA 140312
QY 2878 TTCAATACCTGCCAGCTATTCTTATTTCCACATCCAGCCCTTTGGCTGCTGCTGGTGA 2937
Db 140313 TTCAATACCTGCCAGCTATTCTTATTTCCACATCCAGCCCTTTGGCTGCTGCTGGTGA 140372
QY 2938 AAACATCTCTAGTGTCTTCTCAGCGTTTCCAAAGAAAGATTCCAAATTAACCTTGC 2997
Db 140373 AAACATCTCTAGTGTCTTCTCAGCGTTTCCAAAGAAAGATTCCAAATTAACCTTGC 140429
QY 2998 CAGTCTGAAGAAATCTCCAAACATCCCGACGATCTTGGAGCGCGGCTTTGGGATGG 3057
Db 140430 CAGTCTGAAGAAATCTCCAAACATCCCGACGATCTTGGAGCGCGGCTTTGGGATGG 140489
QY 3058 GACTGCCCGCCGGTCTCTGAACAGGATGCGTGCAGGAGGACACACACACAGCCAGC 3117
Db 140490 GACTGTTTGGCCGGGCTTGAACAGGATTCGTCGCGGTAGG--CACACACACAGCCAGC 140547
QY 3118 CTGTGTGCGGCGGAGTCCGGTCCGGTCCGGTGAAGAGCGCTGCTGGTGGGCGG 3177
Db 140548 C--TGCCTGCGCGCGAGTCCGGTCCGGTGAAGAGCGCTGCTGGTGGGCGG 140605
QY 3178 GGCAGAGCAATTTGTCGAGCGGTACCGAGCCCGCGCTGCCCGGAGGGAGCGG 3237

```
Db 140606 GGCAGAGCCATTGTTCCCGAGCTGACTGAGCCCCCAGCGCTCGCCCCGGGAGGAGCGGG 140665
QY 3238 GCTTCCCGCGTCCCC-AAGTCCAGATCTCTGGGGTGGCTGCCAGTCTCCCTGCCACGCG 3296
Db 140666 GCTTCCCGCGTCCCGAGCTCCCGATCCCGGGTGGCGGCTCGTCTCCCTGCCACGCG 140725
QY 3297 CTTGGGGGACGGAGACGGGACGGAGATGTTAGTGTGGGCGCCCCCGAGGGTTCCAC 3356
Db 140726 CTTGGGGGACGGAGACGGGACGGAGATGTTAGTGTGGGCGCCCCCGAGGGTTCCAC 140785
QY 3357 CACTGTTTCTCTGAGAACTTCCCGAGTGGCCACCCACCCGTTCTCGGTGTGCCGAGGGC 3416
Db 140786 CACTGTTTCTCTGAGAACTTCCCGAGTGGCCACCCAGCGTTCTCGCGTGGCCCCGGGGC 140845
QY 3417 CGGTCTGGGCTAGGCTCGCGGCCCGCCAGCCCCAAAACGGGTGCCAGCGCCCTTCCAGAGA 3476
Db 140846 CGGTCTGGGCTAGACTCGCGGCCCGCCAGCCCCAACACCGGTGCCAGTCCCTTCCAGAGA 140905
QY 3477 GAAAGTCTCCCGAGCGGGATCGCGGCGAGAGGCCACCGCGCGGTGGAAGAGAGCTGAG 3536
Db 140906 GAAAGTCTCTACCGCGGGTTCGCGGCGAGAGGCCACCGCGCGGTGGAAGAGAGCGAG 140965
QY 3537 AAGGAGAAACAGAGGGGAGGGGAGGAGAGTGGCGGCGAGAGGAAACAGCAGATTGCG 3596
Db 140966 AAGGAGAAAGGAGGGGAGGGGAGGAGAGTGGCGGCGAGAGGAAACAGCAGATTGCG 141025
QY 3597 CCGAGCCAATGGCAACGGCAGAGAGGTTGGACCAAAATTCCTTCGGCCAATGACGAGC 3656
Db 141026 CCGAGCCAATGGCAACTGACAGAGAGGTTGGACCAAAATTCCTTCGGCCAATGACGAGC 141085
QY 3657 CGGAGTTTACAGAGCCTCATAGCATTTTCCCGAGAGGCGAGGGGCGAGGGCGCG 3716
Db 141086 CGGAGTTTACAGAGCCTCATAGCATTTTCCCGAGAGGCGAGGGGCGAGGGCGCG 141145
QY 3717 GTGGTGTGTGTGTCGTTGTCGAGCATCCCGCGGCTGCTGCGGTTCGCGCGAGCCTC 3776
Db 141146 GTGGTGTGTGTGTCGTTGTCGAGCAT-CCCGCGGCTGCTGCGGTTCGCGAGCCTT 141204
QY 3777 GGCCTCTGTCTCTCCCTCCCGCTTACCTTCACGCGGAGCCCGCGCGCAGTCAA 3836
Db 141205 GGCCTCTGTCTCTCCCTCCCGCTTACCTTCACGCGGAGCCCGCGCGCAGTCAA 141264
QY 3837 TCTCTCGCATTTGCCCCCTGTTGCGAGCGGATAAAGGGGCTGAGGAATACCGGACA 3896
Db 141265 TCTCGGACATTTGCCCCCTGTTGCGAGCGGATAAAGGGGCTGAGGAATACCGGACA 141324
QY 3897 CGGTACCCGTTGC 3910
Db 141325 CGGTACCCGTTGC 141338

RESULT 8
AC026883
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-799L19 map 3p, WORKING DRAFT
SEQUENCE, 41 unordered pieces.
AC026883.3 GI:8101285
VERSION
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 179879)
AUTHORS
Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
Dong, W., Fan, H., Feng, X., Guan, Q., Gu, X., Guo, D., He, L., Hu, S.,
Huang, F., Jin, Y., Kang, N., Li, C., Li, G., Li, J., Li, L., Li, S.,
Li, T., Liu, Y., Liu, B., Liu, Y., Li, W., Li, W., Li, Y.,
Luo, J., Niu, Y., Qi, Q., Qi, X., Song, S., Sun, M., Sun, W., Sun, Y.,
Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R.,
Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y.,
Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H.,
Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y.,
```

```
TITLE
JOURNAL
REFERENCE
AUTHORS
Zhang, Z., Zhu, B., Yu, J. and Yang, H.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 179879)
Li, W., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W.,
Sun, X., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X.,
Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G.,
Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.
and Yang, H.
Direct Submission
Submitted (25-MAR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On May 29, 2000 this sequence version replaced gi:7417472.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgc@igtp.ac.cn
-----Summary Statistics
Center project name:if project
Center clone name: RP11-799L19
-----Project Information
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149369 bases at least Q40
Consensus quality: 160276 bases at least Q30
Consensus quality: 167926 bases at least Q20
Insert size: 148122; sum-of-contigs
Quality coverage: 3.56x in Q20 bases;sum-of-contigs
```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
* 1
* 1497: contig of 1496 bp in length
* 1596: gap of unknown length
* 1997: contig of 1353 bp in length
* 2949: gap of unknown length
* 3049: gap of unknown length
* 3050: contig of 1360 bp in length
* 4409: gap of unknown length
* 4510: contig of 1628 bp in length
* 6138: gap of unknown length
* 6238: contig of 1047 bp in length
* 7285: gap of unknown length
* 9013: contig of 1629 bp in length
* 9113: gap of unknown length
* 10444: contig of 1331 bp in length
* 10445: gap of unknown length
* 12153: contig of 1609 bp in length
* 12253: gap of unknown length
* 14386: contig of 2133 bp in length
* 14486: gap of unknown length
* 16990: contig of 2504 bp in length
* 17090: gap of unknown length
* 17091: contig of 1932 bp in length
* 19023: gap of unknown length
* 19123: contig of 1375 bp in length
* 20497: gap of unknown length
* 20597: gap of unknown length
* 22309: contig of 1612 bp in length
* 22310: gap of unknown length
* 24677: contig of 2368 bp in length
* 24778: gap of unknown length
* 27655: contig of 2878 bp in length
* 27755: gap of unknown length
```

```

* 27756 30189: contig of 2434 bp in length
* 30190 30289: gap of unknown length
* 30290 33658: contig of 3369 bp in length
* 33659 33758: gap of unknown length
* 33759 36912: contig of 3154 bp in length
* 36913 37012: gap of unknown length
* 37013 41069: contig of 4057 bp in length
* 41070 43658: contig of 2489 bp in length
* 43659 43758: gap of unknown length
* 43759 46574: contig of 2816 bp in length
* 46575 46674: gap of unknown length
* 46675 49890: contig of 3216 bp in length
* 49891 53751: contig of 3761 bp in length
* 53752 53852: contig of 3527 bp in length
* 53853 57378: gap of unknown length
* 57379 61598: contig of 4120 bp in length
* 61599 61698: gap of unknown length
* 61699 65687: contig of 3988 bp in length
* 65688 70673: contig of 4887 bp in length
* 70674 70773: gap of unknown length
* 70774 76442: contig of 5669 bp in length
* 76443 80193: contig of 3651 bp in length
* 80194 80293: gap of unknown length
* 80294 83630: contig of 3337 bp in length
* 83631 83730: gap of unknown length
* 83731 89196: contig of 5466 bp in length
* 89197 93491: gap of unknown length
* 93492 93591: contig of 4195 bp in length
* 93592 99952: contig of 6361 bp in length
* 99953 100052: gap of unknown length
* 100053 107237: contig of 7185 bp in length
* 107238 107337: gap of unknown length
* 107338 113752: contig of 6415 bp in length
* 113753 113852: gap of unknown length
* 113853 119355: contig of 5503 bp in length
* 119356 127315: contig of 7860 bp in length
* 127316 127415: gap of unknown length
* 127416 137358: contig of 9943 bp in length
* 137359 137458: gap of unknown length
* 137459 148311: contig of 10853 bp in length
* 148312 148411: gap of unknown length
* 148412 160062: contig of 11651 bp in length
* 160063 160162: gap of unknown length
* 160163 179879: contig of 19717 bp in length.

```

FEATURES

source

1..179879

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="3"

/map="3p"

/clone="RP11-799L19"

1..1496

/note="assembly_name:Contig7"

1597..2949

/note="assembly_name:Contig9"

3050..4409

/note="assembly_name:Contig10"

4510..6137

/note="assembly_name:Contig11"

6238..7284

/note="assembly_name:Contig12"

clone_end:SP6

vector_side:right

7385..9013

/note="assembly_name:Contig13"

```

misc_feature 9114..10444
              /note="assembly_name:Contig14"
misc_feature 10545..12153
              /note="assembly_name:Contig15"
misc_feature 12254..14386
              /note="assembly_name:Contig16"
misc_feature 14487..16590
              /note="assembly_name:Contig17"
misc_feature 17091..19022
              /note="assembly_name:Contig18"
misc_feature 19123..20497
              /note="assembly_name:Contig19"
misc_feature 20598..22209
              /note="assembly_name:Contig20"
misc_feature 22310..24677
              /note="assembly_name:Contig21"
misc_feature 24778..27655
              /note="assembly_name:Contig22"
misc_feature 27756..30189
              /note="assembly_name:Contig23"
misc_feature 30290..33658
              /note="assembly_name:Contig24"
misc_feature 33759..36912
              /note="assembly_name:Contig25"
misc_feature 37013..41069
              /note="assembly_name:Contig26"
misc_feature 41170..43658
              /note="assembly_name:Contig27"
misc_feature 43759..46574
              /note="assembly_name:Contig28"
misc_feature 46675..49890
              /note="assembly_name:Contig29"
misc_feature 49991..53751
              /note="assembly_name:Contig30"

```

Query Match

Best Local Similarity 37.0%; Score 1448.2; DB 2; Length 179879;

Matches 1456; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 3 GAATTCATCCATTAAATCATACAAATTTAATGGCTTTTAGTATATTCACAGGTTGGCAT 62
DB 109955 GAATTCATCCATTAAATCATACAAATTTAATGGCTTTTAGTATATTCACAGGTTGGCAT 110014
QY 63 CCATCACATCCATTTAGAACAGTTTATTAATCTCAAAAATAAACCCCTGCATTCCTTAG 122
DB 110015 CCATCACATCCATTTAGAACAGTTTATTAATCTCAAAAATAAACCCCTGCATTCCTTAG 110074
QY 123 CCATCACCCCCCAACATCCTCCATCCTCTTCCAGCCCTGGCAACCAACCAATCTACTT 182
DB 110075 CCATCACCCCCCAACATCCTCCATCCTCTTCCAGCCCTGGCAACCAACCAATCTACTT 110134
QY 183 TCTGTCTCTATAAATTTGCCAATTCGGACATTTTCATATAAATGAAGCAACCAACATGT 242
DB 110135 TCTGTCTCTATAAATTTGCCAATTCGGACATTTTCATATAAATGAAGCAACCAACATGT 110194
QY 243 GAGACTTTGTGACTGGCTGCTTTCATTTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA 302
DB 110195 GAGACTTTGTGACTGGCTGCTTTCATTTAGCATTTCTATTTTAAAGGCTCATTTATGTTACA 110254
QY 303 GTACTTAGCAGTACTCTCATTTCTTTTATTTCTCAATGGTATTCACATGCTGGGTATCC 362
DB 110255 GTACTTAGCAGTACTCTCATTTCTTTTATTTCTCAATGGTATTCACATGCTGGGTATCC 110314
QY 363 CATATCATATTTATAGACAGGTTCTCACTCTGTCAACCCAGGCTGGAGTGCAGTGGCAC 422
DB 110315 CATATCATATTTATAGACAGGTTCTCACTCTGTCAACCCAGGCTGGAGTGCAGTGGCAC 110374
QY 423 AATCATAGCTCACGTAACTCAACTCCTGGGCTCAAGTGATCCTTACTACCTCAGGCTC 482
DB 110375 AATCATAGCTCACGTAACTCAACTCCTGGGCTCAAGTGATCCTTACTACCTCAGGCTC 110434
QY 483 CAGAGTAGTACTAGGACTACAGGCAACACAGCCATCTGGCTAATTTTTTTTTTAATTT 542

```



```
QY 2244 GAGGATGAGTGGATGAGTGTCCATCTGAAATCATCCAGTCTTGTTCAGAACTTTCA 2303
Db 120191 GAGTCAATTTCTAGATGAGTGTCCATCTGAAACATCTAGTCCCATTTCA 120250
QY 2304 CACCGGACGAGGAGCAGGACGTGGAAATGAGTCTCTGGTCACTGCGCAGAGAGTTGCC 2363
Db 120251 CACTGGACAGAAAGCCAGGACAGGAATAGACATCTGGTCACTGGCCAGAAAGTTGCC 120310
QY 2364 TTGACCCCTGAGACCACTGGCCAAAGAGAGAGTCTTAACTTACCTCCAGAAAT-CCC 2422
Db 120311 -TGACCCAGAGACAGCCGACAGAGAGTCTTAACTTACCTCCAGAAAT-CCC 2422
QY 2423 AGGTGCTTGTCTTCTGCGAAAGTGAATCATTTGGCGACATCTCGGTATTTTCTCTCTTC 2482
Db 120370 AAGTGTCTCTCTCTGAGAAAGCAGATGGGTGGCACAACTTCTTATTTTGTCTCTTC 120429
QY 2483 CCAGGGAAGAGATCTTAGGAGAGTATTTGGGAAGACATGGGCAATGGAAGACACCGGT 2542
Db 120430 CCAGGAAGAGATCTTAGGAGAGTATTTGGG---GAGACAGCAGGGAAGAGACTGGT 120486
QY 2543 GAATGATAGTCTGCTGGTCTGAGCTCTCATGTAAAGCTCTTACAGACACGGAAG 2602
Db 120487 GAATACAGAGCTGCTGGTCTGAGTCTTACGTGAGGCTCTCCAGGCGCCAGAAAT 120546
QY 2603 ATGGGGGACAGGACAGATCAGTAGGCTCAGAGCATCTCAGGACCCAGGCAATATGG 2662
Db 120547 ATAGGGGCAAGAG-----AGACAGGTCAAGACATCTCAGGACTAGGGCC----- 120591
QY 2663 TCCTGACAGGATTAAGACTTGGGCTCTCATATGTTGTTTCGGCTCAACTGCCAGC 2722
Db 120592 TTCTGATGGGATTAAGAGCGCGCGCTTGCAATTAATTTCTGGG---CAATTCCTAGC 120649
QY 2723 TCCTGATCTTAAATGGAATAGTGGGTTACCCACTCCAGGCTCAGAGAGGCTTA 2840
Db 120710 TTCTCATCTTAAATGGAATAGTGGGTTACCCACTCCAGGCTCAGAGAGGCTTA 120768
QY 2841 CAGAAAGATTTCTGGAATGGCTTGCAGTAAATTAATCAATACCTGCCAGTATCTT 2900
Db 120769 ----AATGAATCTGGAATGGCTTGCAGTAAATTAATCAATATCTGTAATCTT 120824
QY 2901 ATTCACATCCAGCCCTTTGCGCTGCTGCTGGTGAAGACATCTCAAGTCTTCTGA 2960
Db 120825 -TTCCACCTAGCAGTCTTTGCTGATCTGGTGAAGACATCTGCTTCTGCTTCTGCT 120883
QY 2961 CGGTTTCCAGAAAGATTTCCAAATTAACCTG-CCAGTCTGAAGATCTTCCAAAC 3019
Db 120884 TTTTTCATTAAGA--GTTCCAAATTAACCTGACAAATCTGAAGATCTTCCAAAC 120941
QY 3020 ATCCCGCAGCATCTTGGAGGCGCGGCTTGGGATGAGTCTCCCGCGGCTGTA 3079
Db 120942 GTCCCGCAGCGTTCTAGAGCGTGGGCGAGCGGCTGCTGCGCCACTGGTGGCG 121001
QY 3080 CAGGA-----TGCGTGGCGCGCAGGACACAC 3105
Db 121002 GTGAGCTGCCCTGCGAGCGTACCCAGTGGGCGCACAGTGGCGCGCAGGACAC 121061
QY 3106 ACACAGCAGCCTGTGTGTG-----CGGCGGAGTCCGGTGGCTCCCGGTGAGCAG 3159
Db 121062 GTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121120
QY 3160 CGCGTGGCTGTGTGGGCGGCGAGGACATTTGTCAGGCGTACCGAGCCCCCGGCTC 3219
Db 121121 CGCGTGGCTGTGTGGGCGGCGAGGACATTTGTCAGGCGTACCGAGCCCCCGGCTC 121180
QY 3220 GCCCGGAGGAGGCGGCGCTTCCCGCGTCCCAAGCTCCAGATCTCGGGTGGCTGCA 3279
Db 121181 GCGGACAGGAGGCGGCGCTTCCCGCGTCCCAAGCTCCAGATCTCGGGTGGCTGCA 121239
QY 3280 CGTCTCTCTGCGAGCGGCTGGGCGGAGCGGGAAGACGGGACGGAGATTTAGTGTGGC 3339
```

```
Db 121240 GTCTGCCCCTCCCGCGCGCTTGGGCGGACGCGAAATGCGGGAAGAGCTTGTGAGGAGTGGC 121299
QY 3340 GCCCGCGGAGGTTTCCACCACTGTTTCTGAGAAATTTCCCGAGTCCCGCAGCCACCGGTTCC 3399
Db 121300 G-CAACCGTGGCTGCCCACTGTTTCTGTGAAATTTGCCAGTGGCCACCA-----TTC 121354
QY 3400 TCCGTGTGCCGAGGCGGCTCTGGGCTAGGCTCCGCGCCCGCAGCCCAACCGGTTCC 3459
Db 121355 GCCCTGTGCCCGGCGC---ACCCTGACTTGGAGCACCCCGCTTCCGAACTCGATCC 121411
QY 3460 CCAGCCCTTCCAGAGAAAGTCTCCGACGCGGATGCGCGGAGAGGCGCCAGCGCGG 3519
Db 121412 GTCCTCTCTCACACAGGAGAGCTCCCGA-----GGCAGGCGCCAGCGCGC 121458
QY 3520 GTGGAAGAGAGCTGAGAAGGAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3579
Db 121459 GGGGAAGAGAGCGC---AGGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121515
QY 3580 GGGAAACAGCAGATTGCGCCGAGCAATGGCAACGCGAGGACGAGTGGCACCAATTTCC 3639
Db 121516 GGGAAACAGCAGATTGCGCCGAGCAATGGCAACGCGAGGAGGAGTGGCACCAATTTCC 121575
QY 3640 TTGGGCAATGACGAGCGGAGTTTACAGAGCTTANTAGCATTTTCCCGAGGAGGAGG 3699
Db 121576 TTGGGCAATGACGAGCGGAGTTTACAGAGCTTANTAGCATTTTCCCGAGGAGGAGG 121635
QY 3700 GCAGGCGCAGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3759
Db 121636 GCAGGCGCAGCGC-----GGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121690
QY 3760 CGGTGCGCGGAGCGCTCGG-CTGTGTCTCTCTCC-CTCCCGCGCTTACTTCCAGCGG 3817
Db 121691 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121750
QY 3818 GACGCGCGCGCGGAGTCACTCTCGCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3877
Db 121751 GACTGTCTGGCGGAGTCACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121810
QY 3878 GTTGAAGAAATCCGACACGCTACCCCTTGC 3910
Db 121811 TCTGAGAAATCCGACACGCTACCCCTTGC 121843
AY241932 17088 bp DNA linear MAM 02-APR-2003
Bos taurus stearoyl-CoA desaturase variant A (SCD) gene, complete cds
ACCESSION AY241932
VERSION AY241932.1 GI:29469126
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
REFERENCE 1. (bases 1 to 17088)
AUTHORS Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
TITLE Genomic sequence of the bovine stearoyl-CoA desaturase
JOURNAL Unpublished
REFERENCE 2. (bases 1 to 17088)
AUTHORS Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) Animal Science, University of California, One Shields Avenue, Davis, CA 95616-8521, USA
FEATURES
source
1. .17088
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/chromosome="26"
1. .17088
gene
```


Db 635 CTTACTCTCACAGGGCCACAAGTTAAGATCAAAATGAATCTCTGTGGATTGCTTG- 693
QY 2870 AGTAATAATTCAATACCTGCCAGCTTCTTATTCATCCATCCAGCCCTTTCGCTGCTG 2929
Db 694 -----CAATACCTTTGGCTACTCTTCTGCTCACTTAAGTACTTTCTCTGCTT 743
QY 2930 CTGGGTGAAAACACATGTCAGTGTTC-----TGACGGTTTCCCAAGAAAGATTC 2981
Db 744 CTGAGTGAAGACGCTCTCTCTTCTCCAGAAATGACCTTTCCCAAGAAAGCAATTC 803
QY 2982 CAAATATACACCTGCCAGTCTGAAGATCTCCAAAACATCCCGACGCAATCTTGGAGGC 3041
Db 804 CAAACACAGCCCGAGCTTCT-----AATCTCCAAAACATCCCGACGCAATCTTGGC 859
QY 3042 CGGGCTTTGGGATGGGACTGCCCGCCCGGCTC-----TGCGGTTC----- 3074
Db 860 ATGGGCACAGGACAGACACACTCTGCTGCTGGCGGTGAACATTTCTTCCGGGTG 919
QY 3075 -----CTGAACAGGATGGTGGCGGACAGCACACACAGCAGCCTGTGTG 3126
Db 920 ACCATTGGCTGAACCTCTTACGGCGGTGCAGATCTACATCAGCGCATCTACTGTG 979
QY 3127 CGGCGGAGTCCCG-----TGCGGTCCCG 3150
Db 980 TGGCGGCTTCCAGCTGCTGCTGCGGAGGGGAGACCCGCAACCTGTTTCCAGCAGTCCCG 1039
QY 3151 GTGAGCAGCGCTGCTGCTGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3206
Db 1040 GGTAGCGCTGCGCGCTGCTGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1099
QY 3207 -----GCGCGCGCTGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3244
Db 1100 TCCCTGCGCGCTTACCGCGCATCCCGGCGCTGCGGACAGAGGAGGAGGAGGAGGAGG 1159
QY 3245 GCGTCCCGCAAGCTCCAGATCTGCGGTGCTGCGAGTCTCCCTGCGGAGGAGGAGGAGG 3304
Db 1160 CGGAGCGCGGCGCGGCTGCTGCGAGTCTGCGAGTCTGCGGAGTCTGCGGAGTCTGCG 1219
QY 3305 GACGGGAAGACGGGAGGAGTGTAGTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1279
Db 1220 AGCGGGAAGGCTGGGTGGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGT 1334
QY 3365 CTTGAGAACTTCCCGAGTCCACCCCGCTTCTGCGTGTGCGGAGGAGGAGGAGGAGGAGG 3424
Db 1280 CCGAGAACTTCCCGAGTCCACCCCGCTTCTGCGGAGTGTGCGGAGTGTGAGTGTGTG 1394
QY 3425 GCGTAGGC-----TCCG 3480
Db 1335 GCGTGGCAGCATCCCGCGCGCATCTCCCGAGTGGGTCTCTCCCTCTCCCGCGCGCGCT 1394
QY 3481 GCTCCGCGCGGATGCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3540
Db 1395 CAGAGCG--GAGAGGTGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3600
QY 3541 AGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
Db 1450 AGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
QY 3601 GCCAATGCAACGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
Db 1510 GCCAATGCAACGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1569
QY 3661 GTTTACAGAGCTTATAGCATTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3720
Db 1570 GTCTACAGAGGCG--CATTAGCATTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1628
QY 3721 TGTGCTGCTGCTGCGCAGCATCCCGCGCGCGCTGCTGCGGTGCGCGGAGGAGGAGGAG 3780
Db 1629 GCGAGCGCGGCTGTGTGCTGAGCATCAGTCTTCTGCTTCTTCTGCGCGCGCGCGCG 1686
QY 3781 TCTGTCTCTCTCCCGCTCCCGCGCTTACCTCCAGCGGAGGAGGAGGAGGAGGAGGAGG 3840

Db 1687 TCGCGGCT 1746
QY 3841 TCGCACATTTGCCCTGCTTGGCAGCGGATAAAGGGGCTGAGGAAATACCGGACACGGT 3900
Db 1747 GCGCACATTTGCCCTTGTGGCAACGAATAAAGAGGTCTGAGGAAATACCGGACACAGT 1806
QY 3901 CACCGGTGC 3910
Db 1807 CACCCCTGC 1816

RESULT 11

AL157935

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL157935 164201 bp DNA linear PRI 09-JAN-2002
Human DNA sequence from clone RP11-203J24 on chromosome 9, complete
sequence.
AL157935
AL157935.28 GI:17221172
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164201)
Corby, N.
Direct Submission
Submitted (09-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Nov 30, 2001 this sequence version replaced gi:16944853.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-203J24 is from the library RP11-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VRCOR: pBACe3.6

This sequence is the entire insert of clone RP11-203J24 The true
left end of clone RP11-379C10 is at 163899 in this sequence. The
true right end of clone RP11-228B15 is at 66809 in this sequence.
FEATURES
Location/Qualifiers
1..164201
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="9"
/clone="RP11-203J24"
/clone_lib="RP11-11.1"
55900
/note="Tandem repeat. Forced join. Gap size estimated to
be approximately 280bp by restriction digest data."
62058..62326
/note="Sequence from overlapping clone RP11-228B15.
Assembly confirmed by restriction digest data."

source

misc_feature

misc_feature

misc_feature 153662..153692
/note="single clone region. Assembly confirmed by
restriction digest data."

ORIGIN

Query Match 11.8%; Score 460; DB 9; Length 164201;
Best Local Similarity 58.9%; Pred. No. 3.3e-101;
Matches 1107; Conservative 0; Mismatches 660; Indels 111; Gaps 14;

365 TATCATATTATTAGACAGAGCTCTCACTCTGTCACCCAGGCTGAGTGGACAA 424
5030 TTTTCTTTCTTTTGGACAGAGCTTGTCTGTACCCAGGCTGGAACCCAGTGGCA 5089

425 TCATAGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGTACTTACCTCAGCCCTCCA 484
5090 TCTCAGCTCACTACAACTCTGTGCTCTGGGTTCAGCAATTCCTCTGCTCAGCCCTCT 5149

485 GAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 544
5150 GAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 5209

545 ATTTTGTATGATCAATTTCT-----TCTTTTGTGTGTGTGTGTGTGTGTGTGT 595
5210 GAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 5269

596 GTCTCACTTTGTACCCAGGCTGGAGGCTAGTGTGATGTGACAGCTGA--GCAGCCTTG 653
5270 GTCTCACTTTTCCAGGCTAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 5329

654 ACTTCTGGGCTCAAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 713
5330 ACCTCTAGGCTCAAGCAATCTCCCACTCGGCTTCCCAAGTGTGGAATACAGACA 5389

714 CGTGTACCATGC--CTGGCTGATATTTTCTTGTGAACAGGCTATCAGTCTGTGTC 770
5390 TGAGCTACATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 5449

771 CCAGGCTGGAGTACAGTGGGCTAATAAGTCACTGCAGCCTCCCTCTCGGCTCAAG 830
5450 CCAGGCTGGATTCAGTGGGACACCTCGCTCACTGCACACCTCTCGCTCTGGGCTCAAG 5509

831 CAATCCGCTGGCTCAGCACTCTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 890
5510 CAATCTCTGCTCAGCACTCTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 5569

891 GCTAAGTTTAAATGATTTTGTGTATAGAGAGTCTGTGTATGTGCTCAGGCTGT 950
5570 GCTAATTTT-----GTATTTTGTAGATGGGGTTTCAGCATGTGTGGCCAGACTGG 5623

951 ATTTTATTTGTGAGCAAGGCTCACTATGTGTCCATGATCCCGCCACCTCCACCTCCC 1010
5624 TCTCAAACTCTTGA-----CCTCAATGATCTGCCCGCTCGGCTCCC 5667

1011 AAGTGTCTCATCTATCTCTTATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 1070
5668 AAGTGTCTGGAATTCAGGCTAGGCTCAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 5727

1071 TTGACCATTTATGAATAATCTCAGTGAATATTCATGTATATATTTGTGTGGCATATGT 1130
5728 CAGATGAGGGAACAGACCTTGAGAGTGAAGTCAATTTGCTTGTATGACACAGGAAGT 5787

1131 TTTTATTTCTGTGGGTTTATATCTAGGAGTGGAAATTTGTGGATCCCGGGTAAATTTTG 1190
5788 CAGCCAGGGGCTGGGATTCGTGCCAGGTAGGCTGTCTCTCACATTTTCTTTCTTT 5847

1191 ACAGGCAAGTTTCAGGGGAAGAAACCTTGGGAAATGAAGCATGTTTGAATCAGCAA 1250
5848 TCACCTTTTCTTTTCTGTGAGATGGAGTTTCACTCTT---CCTGCCAGGCTGGAGTGC 5904

1251 GAGTGTAGGAGGTTTTCGGAGTTTATTTTATATCTGTGTGCAATATGTCAGTTTGTATG 1310
5905 GAGACAGGCTCATCAACATGTGTGAACCCCAATGGTGGATCTTAGTCTCACTGTCTTT 5964

1311 AAGATCAAGTTTACTTAAGTGTGAAGTGTGAATTAAGCTGTGAATAGGGGCTTCAGAGT 1370
5965 TTTTGTAGGCTCGT--CAAGTGAAGCAATGAGAGTGGAGGAAGCAAAAGCAATCTGTAA 6022

1371 AAAATCATGAAGCACACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATA 1430
6023 CTAGTGTGATCAATTAATCTTGTAAACACAATGTGTGTTCATTTTATGTGTGCCATTAG 6082

1431 AAAATCAAAATTTTGT 1490
6083 TTTTGTCTATTTATTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6141

1491 GGAGT 1550
6142 TGCATGGCCGATCTCGGCTCACTGCACTCCCTCCCTCTGTGTGTGTGTGTGTGTGTGTGT 6201

1551 GCTTCAGCCTCCCAAGTAGT 1610
6202 GCCTTAGCCTCCCAAGTAGT 6261

1611 GTATTTTGT 1663
6262 GTATTTTGT 6321

1664 ----TTTGT 1719
6322 TCAGTGTATCCACCCGCTTGGCTCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6381

1720 ACCAAAAAATAAATCTTTAGTAATCCAAATCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1768
6382 ACCCGCGCTCTCAATTTTCCACACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6441

1769 CTGTCTAGT 1828
6442 TTTTCTAACTTATTTCCCAAGAAAGAAAGAACACAGCAATCTACAAGCCAGTGTCTATT 6501

1829 AAGAGAAATAGTAATTAATTTCT-----TTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1853
6502 GAAAGAAATCACTTCAAGACTTTTCTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6561

1854 -----CCGGCCAGGCTGT 1905
6562 TAAGACAGGT 6621

1906 AGCATGT 1965
6622 AGGTGAGCCGATCACCTGT 6681

1966 CCGTCTCTACTAAATAACAAATAATAGCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2025
6682 CCGTCTCTACTAAATACTA--AATTAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6739

2026 GTTACTTGGAGGT 2084
6740 GCTACTCGGAGGCTGT 6799

2085 AGCCAGATCACACCTGT 2143
6800 AGCCAGATCGCGCATTTGAGCTCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6859

2144 AAACAAAAACAAAAAA 2161
6860 AAAAAAATAAAAAAA 6877

RESULT 12
AX357255
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX357255
Sequence 1 from Patent WO0162954.
AX357255
AX357255.1 GI:18674409
Homo sapiens (human)

617 bp
DNA
linear
PAT 13-FEB-2002

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Brownlie,A.J., Hayden,M.R., Attie,A.D., Ntambi,J.M.,
Gray-Keller,M.P. and Miyazaki,M.
TITLE Methods and compositions using stearyl-coa desaturase to identify
patent: WO 0162954-A 1 30-AUG-2001;
JOURNAL Xenon Genetics Inc. (CA) : WISCONSIN ALUMNI RESEARCH FOUNDATION
(US) : University of British Columbia (CA)
FEATURES
source 1..617
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 11.5%; Score 450.4; DB 6; Length 617;
Best Local Similarity 99.8%; Pred. No. 3e-99;
Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3459 CCCAGCCCTTCCAGAGAGAAAGCTCCGACGCGGGATGCGGCGAGAGCCAGCGCGG 3518
Db 4 CCCGCCCCCTTCCAGAGAGAAAGCTCCGACGCGGGATGCGGCGAGAGCCAGCGCGG 63
QY 3519 GTTGGAGAGAGCTGAGAGAGGAGAAACACAGAGGGAGGGGAGCGAGCTGGCGGCGAG 3578
Db 64 GTTGGAGAGAGCTGAGAGAGGAGAAACACAGAGGGAGGGGAGCGAGCTGGCGGCGAG 123
QY 3579 AGGGAACAGCAGATTGCGCGAGCCATGCGACGCGGAGAGCAGAGTGGACCAATTTC 3638
Db 124 AGGGAACAGCAGATTGCGCGAGCCATGCGACGCGGAGAGCAGAGTGGACCAATTTC 183
QY 3639 CTTGCGCCAAATGACGAGCGCGAGTTTACAGAGCCCTATTAGCATTTCCCGAGAGCAGG 3698
Db 184 CTTGCGCCAAATGACGAGCGCGAGTTTACAGAGCCCTATTAGCATTTCCCGAGAGCAGG 243
QY 3699 GCGAGGGGACAGCGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3758
Db 244 GCGAGGGGACAGCGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 303
QY 3759 GCGGTGCGCGCAGCGCTCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3818
Db 304 GCGGTGCGCGCAGCGCTCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363
QY 3819 ACCGCGCGCGCAGTCAACTCTCGACATTGCGCCCTGCTTGGCAGCGGATAAAGGGGG 3878
Db 364 ACCGCGCGCGCAGTCAACTCTCGACATTGCGCCCTGCTTGGCAGCGGATAAAGGGGG 423
QY 3879 CTGAGGAATACCGACAGCGGTTCACCGGTGCG 3910
Db 424 CTGAGGAATACCGACAGCGGTTCACCGGTGCG 455

```

```

RESULT 13
AC119725
LOCUS AC119725
DEFINITION Homo sapiens chromosome 3 clone RP11-759D18, WORKING DRAFT
SEQUENCE, 36 unordered pieces.
ACCESSION AC119725
VERSION AC119725
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200288)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,p., Burkett,C., Burrell,K.L., Byrd,N.C.,

```

```

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabibi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsdon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegad,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Ogih,M., Okuwonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pichens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Schreier,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodtgren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 200288)
Worley,K.C.
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 200288)
Worley,K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 27, 2003 this sequence version replaced gi:20376806.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HCIR
Center clone name: RP11-759D18
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 169013 bases at least Q40
Consensus quality: 174569 bases at least Q30
Consensus quality: 178019 bases at least Q20
Estimated insert size: 177488; sum-of-coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases
Quality coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases; sum-of-coverage: 3x in Q20 bases
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces

```


* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 1266: contig of 1266 bp in length
1267 1366: gap of unknown length
1367 2448: contig of 1082 bp in length
2449 2548: gap of unknown length
2549 4153: contig of 1605 bp in length
4154 4253: gap of unknown length
4254 5318: contig of 1065 bp in length
5319 5418: gap of unknown length
5419 6987: contig of 1569 bp in length
6988 7088: gap of unknown length
7089 8369: contig of 1282 bp in length
8370 8469: gap of unknown length
8470 9488: contig of 1019 bp in length
9489 9588: gap of unknown length
9589 11671: contig of 2083 bp in length
11672 11771: gap of unknown length
11772 13295: contig of 1524 bp in length
13296 13395: gap of unknown length
13396 14422: contig of 1027 bp in length
14423 14522: gap of unknown length
14523 15982: contig of 1460 bp in length
15983 16082: gap of unknown length
16083 18344: contig of 2262 bp in length
18345 18444: gap of unknown length
18445 20933: contig of 1649 bp in length
20934 21347: contig of 1154 bp in length
21348 21447: gap of unknown length
21448 24953: contig of 3506 bp in length
24954 25053: gap of unknown length
25054 27672: contig of 2619 bp in length
27673 27772: gap of unknown length
27773 29866: contig of 2094 bp in length
29867 29966: gap of unknown length
29967 31985: contig of 2019 bp in length
31986 32085: gap of unknown length
32086 34335: contig of 2249 bp in length
34336 34434: gap of unknown length
34435 37766: contig of 3332 bp in length
37767 37866: gap of unknown length
37867 40539: contig of 2673 bp in length
40540 4375: contig of 3736 bp in length
4376 4475: gap of unknown length
4476 50326: contig of 5851 bp in length
50327 50426: gap of unknown length
50427 53739: contig of 3313 bp in length
53740 53839: gap of unknown length
53840 61829: contig of 7990 bp in length
61830 61929: gap of unknown length
61930 68906: contig of 6977 bp in length
68907 78680: contig of 9674 bp in length
78681 84331: contig of 10651 bp in length
84332 89531: gap of unknown length
89532 100731: contig of 11199 bp in length
100732 100830: gap of unknown length
100831 112190: contig of 11360 bp in length
112191 112290: gap of unknown length
112291 122608: contig of 10316 bp in length
122609 122708: gap of unknown length
122709 135164: contig of 12458 bp in length
135165 135264: gap of unknown length
135265 147115: contig of 11851 bp in length
147116 159369: contig of 12154 bp in length
159370 159469: gap of unknown length
159470 173326: contig of 13857 bp in length
173327 173426: gap of unknown length
173427 200288: contig of 26862 bp in length.
```

FEATURES

```
Location/Qualifiers
1..200288
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="3"
/clone="RP11-759D18"
```

ORIGIN

```
Query Match 11.4%; Score 447.2; DB 2; Length 200288;
Best Local Similarity 58.5%; Pred. No. 4.6e-98;
Matches 1099; Conservative 0; Mismatches 668; Indels 111; Gaps 14;

QY 365 TATCATATTATTAGACAGAGTTCTCAGCTCTGTCAACCCAGGCTGGAGTGCAGTGGCACAA 424
DB 103464 TTTCTTCTTTTGGACAGAGGTCTTGCTCTGTCAACCCAGGCTGGAGTGCAGTGGCACAA 103523
QY 425 TCATAGCTCAGTGTAACTTCAAACTCTCTGGGCTCAAGTGATCCTACTACTCAGCCTCCA 484
DB 103524 TCTCAGGTCACTACAACCTCTGCTTCTGGGTTCAAGCAATTTCTCTGCTCAGCCTCCT 103583
QY 485 GAGTAGCTAGGACTACAGGCACACAGCCAPACCTGGCTAAATTTTTTTTTTAAATTTTC 544
DB 103584 GAGTAGCTAGGACTACAGGCCTGTCCACACACAGATTAATTTTTTATAATTTTAAATA 103643
QY 545 ATTTTATGTATTCATTTTCT-----TTCTTTTGTGTTGTTGTTGTTGAGATAGG 595
DB 103644 GAGATGGGGTTTCACTATGTTGGCCAGGCTGGCTTTCTTTGTTTGTGTTGTTGAGACAGG 103703
QY 596 GTCTCAGTTTGTATCCAGGCTGGAGGTCAGTGGATGGTGCAGCTGA--GCAGCCTTG 653
DB 103704 GTCTTACTCTCTTTTCCAGGCTAGAGTGCAGTGGTGCAATCATGCTCAGTGCACCTCA 103763
QY 654 ACTTCTCGGCTCAAGTATCTCTGCTCAGCTCCCAAGTACCTGGGAGTACAAACA 713
DB 103764 ACCTCTAGGCTCAAGCAATCTCCCACTCGGCTTCCCAAGTCTGGAATACAGACA 103823
QY 714 CGTGTCACCATGC---CTGGCTGATATTTTCTTCTTCAACAGGATCATCTGTGTTGC 770
DB 103824 TGAGCTACCATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 103883
QY 771 CCAGCTGGAGTACAGTGGCTGAATATAGTCTCACTGCAGCTCCCTCTCTGGGCTCAAG 830
DB 103884 CCAGCTGGATTCAGTGGCACAACTCGGCTCACTGCAACCTCTGCTCTCTGGTCAAG 103943
QY 831 CAATCCGCTGCCTCAGCATCTCAGTAGCTGGGACTACAGGCTTGCCACAGGCCCA 890
DB 103944 CAATCTCGCTCAGCATCTCAGTAGCTGGGACTACAGGCGATGCCCACCATGCTG 104003
QY 891 GCTAAGTTTAAAAAATGATTTTGGTATAGAGAGGTCTTGCTATGTTGCTCAGGCTGT 950
DB 104004 GCTAATTTT-----GTATTTTATAGATGGGGTTTCAGCATGTTGCCAGACTGG 104057
QY 951 ATTTTATTTGTTGAGACAAAGTCTCACTATGTTGCCATGATCCGCCACCTCCACTTCCC 1010
DB 104058 TCTCAAACTCTTGA-----CCTCAAAATGATCTGCCGCTCGGCTCCC 104101
QY 1011 AAGTGTCTCATCTTATCTGTTTCATTAGTCAAGTGCAGACATTTAGTGTGTTTCCACTTT 1070
DB 104102 AAGTGTCTGGAATTTACAGGCATGAGCCACTGCACCCGCGCCACTATTCTTATTTTCACTTA 104161
QY 1071 TTGACCATTTAATAATACTCTCCAGTGAATATTCACTATATACATTTGTGTGGGCATATGT 1130
DB 104162 CAGATGAGGGAAACCAAGACCTCGAGGTGAGTCACTTTGCCCTTTGATGACACAGGAAGT 104221
QY 1131 TTTCAATTTCTGTTGGGTTTATATCTTAGGAGTGGAAATTTGCTGGATCCCGGGTAATATTTTG 1190
DB 104222 CAGCCAGGGGCTGGGATTCGTGCCAGGTAGGCTGTGCTCTCACATTTTCTTTCTTTT 104281
```

QY	1191	AACGGCAGAGTTTCAGGGGAGAGAAAACTTGGAAAAATGACATGTTTAGAAATCAGCAA	1250
Db	104282	TCACGTTTTTTTTTTTTCTGAGATGGAGTTTCACTCTT---CGGCCAAGGCTGAGATGC	104338
QY	1251	GAGTCAGAGGGGTTTTTCGAGAGTTTATTTATTTATTTCTGTACAAATGTGCAGTTGATG	1310
Db	104339	GAGACCAAGCTTCATCAATGTGTAACCCCAATGGTGCAATCTTACTCACTGTTCTT	104398
QY	1311	AAGATACAGTTTATCTAAGTAGAGAAATGAAATTAAGCTGGAATATGGGCGTTTCAGAGT	1370
Db	104399	TTTTTTAGCGTCGT--CAAGGAGACATGAGAGTGGAGAAAGAACAAATCTGTAA	104456
QY	1371	AAATCATGAGCACTTTGAAATACCAAAATTAAGAGCGTTGGCTGTAAACAAATATATA	1430
Db	104457	CTAGTTGTGATCAATCTGTGTAACCAATGTTTTTTCATTTTAGTGTCATTG	104516
QY	1431	AAAAATCAAAATTTTTTTTTTTTTTTTATAGAAAGAGCTGCTCTTTCACCTGGCTGGAG	1490
Db	104517	TTTTTTCTATTATATATTTTTTTTGAGATGAGATCTGTGTG--TGCAGGGCTGGAA	104575
QY	1491	GGCAGTGTGTGATCTCAGCTCACTGCACATTTTGGCTCCGGGTCAAGCAATTCCT	1550
Db	104576	TGCATATGGCGAATCTGGCTCACTGCACCTCACTCTGGGTTCAAGTATCTCCT	104635
QY	1551	GCTTCAAGCTCCCAAGTAGCTGGGACATACAGAGACTTCCCACTATGCCAGCTGATTTT	1610
Db	104636	GCCTTGGCTCCCAAGTAGCTGGGACATACAGAGACTTCCCACTATGCCAGCTGATTTT	104695
QY	1611	GTTTATTTAGTAGAGATGGGATTTCACTTTGTGTGGCAAGCTGATCTCAACT-----	1663
Db	104696	GTAATTTTATGTAAGATGGGGTTTGCCATGTTTGGCAGGAATGGGCTCAACTCCTGACC	104755
QY	1664	----TTTGTGTGATTAATGTTGTAACATGTTGTTCTTTGCTGAGGTAGGGCCCCAG	1719
Db	104756	TCAGGTGATCCACCGCGCTTGGCTCCCAAGTGTAGAAATTAATGAGGTAGCCATGCG	104815
QY	1720	ACCAAAAAAATAATCTTAGAATCCAAATCAGTGTG-----GTTGACCA	1768
Db	104816	ACCCGGCGGCTGCACATTTTCCACACATTTTGTGAGCACATCTCTCCTAGACAT	104875
QY	1769	CTGTCACTTGAGAACACAGTGTGACAGAGGCTCAGAGTAGAGTATCTGTGTGA	1828
Db	104876	TTTTCTAACCTTATTTCCAGAGAAAAAAGAACACAGCAATCTACAGCCAGTGTATTT	104935
QY	1829	AAGAGAAATAGATGAAATATATCT-----	1893
Db	104936	GAAAGAAATGACCTCAAAACCTTTCTGAACTATGTGCAATGTGAACAGGTGAGTGTGAGAC	104995
QY	1854	-----CGGGCCAGGCGTGTGGTGCATGCTGTATATCCGACCTTTGGGAGGCCA	1905
Db	104996	TAAGACAGGTGAGACAGGCGTGTGGTGTACGCTGTATATCCAGACCTTTGGGAGGCTG	105055
QY	1906	AGGCATGTGATGACTGTGAGGTCAAGAGTTCAAAACCAAGCTTGCCCAACATGTTGAACCT	1965
Db	105056	AGGTGAGCCGATCACTGTAGGTTGGGAGTTCAAGACCAACCTTACCAACATGTGAGAAACC	105115
QY	1966	CCGTCTCTAATAAAATACAAAAATTAAGCTTAAGTGTGTGGCGCATGCTGTAAATCCCA	2025
Db	105116	CCGTCTCTAATAAAATATA--AATTAGCAGAGTGTGTGGCGCATGCTGTAAATCCCA	105173
QY	2026	GCTACTTGGGAGGGTGGAGAGAGAAATTTCTTGA--CCCGGAGGCAAGGTTTGCAGTG	2084
Db	105174	GCTACTCTTTTGTGTGGCGAGAGATGCTTGAACCCCATAGGCTTAGGTTGTGGTG	105233
QY	2085	AAGCAGATACACCACTGCACTCCAGCGTGGG--GGAGAGAGGAGAACTTCTCTCAAAA	2143
Db	105234	AGCCAAAGATGCGCCCATGTGAGCTCTAGCTGTGGCGACAGAGTGAATCTCGCCTCAAAA	105293
QY	2144	AAAAAATAAAACAAAGAA	2161
Db	105294	AAAAAAAAAAAAAAAAAAAA	105311

RESULT 14	
AL731547	
LOCUS	AL731547
DEFINITION	AL731547 175842 bp DNA linear PRI 27-SEP-2002
ACCESSION	Human DNA sequence from clone Rp11-565H3 on chromosome 10, complete sequence.
VERSION	AL731547 AC073586
KEYWORDS	AL731547.9 GI:23380932
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 175842)
AUTHORS	Wray, P.
TITLE	Direct Submission
JOURNAL	Submitted (27-SEP-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humanyes@sanger.ac.uk On Sep 30, 2002 this sequence version replaced at: 2160432
COMMENT	

Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver
 Street, Waltham, MA 02453, USA
 http://www.genomecorp.com
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: hunquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

These sequences were annotated as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, EMBL; SWISSPROT, Tr.; TREMBL, Wp.; WORMPEP, information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/c_elegans/wormpep. This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/chr10/>.

RP11-565H13 is from the library RPci-11.2 constructed by the group of Pleier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6.

```
FEATURES
SOURCE
Location/Qualifiers
1..175842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-565H13"
/clone_1b="PFC1-11.2"
```

Query Match	11.1%	Score 433;	DB 9;	Length 175842;
Best Local Similarity	58.3%;	Pred. No. 1.3e-94;		
Matches 1071; Conservative	0;	Mismatches 670;	Indels 96;	Gaps 14;

QY	365	TATCATATTATTAGACAGAGTTTCACCTGTGACCCAGGCTGAGT---	GCATGGCA	421
Db	146700	TTTTTTTTTTTTTGGAGACGATCTTGCTGTGTACCCAGGCTGAGCTCCGAGAGGCA	146759	
QY	422	CATCATAGCTACCTGTAACTCCTCTGGGCTGAAGTGAATCTTACTACGACCT		481

[illegible]

Db	147780	CTCAGGCTGGAGTGC	CAATGSCAATCTCAGTTCACTGCAACTCTGCTCCAGGTTCA	147839
QY	1539	AGCAATTCCTCGCTT	CAGCCTCCCAAGTAGCTGGGACTACAGGCACITTC	147899
Db	147840	AGTGAATTCCTCGCT	CAGCCTCCCTGAGTAGCTGGGATTTACAGGGGTGTGCACCAACAC	147899
QY	1599	CAGCTGAATTTTGTAT	TTTATTAGTAGAGATGGGATTTCACTTTGTGTGGCCAAAGCTGTCTC	147959
Db	147900	CGGCTAAATTTTGTAT	TTTATTAGTAGAGACGGGGCTTCACTAGTTGGTTCAGCTGTCTC	147959
QY	1659	AAACTTTTGTCTGCT	ATATTTCTTGACTATTTCTTTTGTCTG-----AGTGGGGC	1714
Db	147960	AAACTTTTGAATCT	CAGGCGATCCACCCACTCCCAAAATGCTACAAATTTATAGC	148019
QY	1715	CCAGACCAAAAAAATA	ATTTAGAAATCCAAATCAGTGTGTGGTTGACCACTGTCA	1774
Db	148020	GTGAGCCACTG	CACCTGGCCAAATATTTATCTTATGTTGGAATATTTCAAGTC	148079
QY	1775	CTTGAGAACCA	CAGTGTGACAGGCGCTCAGGAGTAGAGGTGATC-----TCTGTCTCGAAA	1830
Db	148080	CTCTCTTCTATCT	GCTTTTGAATATATCTATGCTGGCTTAGTGGCTCCACT	148139
QY	1831	GAGAAATAGAAT	GAAAAATATTC-TCCGGGCCAGGGCTGTGGCTCATCGCTGTATATGCT	148199
Db	148140	ATCAAAACAT	CATAATTTGTTTCTTAGCCAGGGGTGTGGCTCATCGCTGTATATGCT	148199
QY	1890	GCATTTGGGAGGC	CAAGGCATGTGGATCACTCAGGTTCAGGAGTTCAAAACCAAGCTGG	1949
Db	148200	GCATTTGGGAGGC	CAAGGCAGAGTCACTCAGGTTCGGAGTTTCAGACCAAGCTGA	148259
QY	1950	CCAACTGGTCAAA	CCCCCGTCTCTACTAAAAATCAAAAAATTAAGTGTGGTGGC	2009
Db	148260	CCAATGGAGAA	CCCCCGTCTCTACTAAAAATCAAAAAATTAAGTGTGGTGGC	148319
QY	2010	CATGCTGTAT	CCTCAGTCTCTGGGAGGTGAGCGAGGAAATTTCTTGAACCCGGAG	2069
Db	148320	CATGCTGTAT	CCTCAGTCTCTGGGAGGTGAGCGAGGAAATTTCTTGAACCCGGAG	148379
QY	2070	GCAGAGTTG	CAGTGAAGCGAGATCAACCACTGCATCCAGCTGGGGGAGAGAGCGA	2128
Db	148380	GCAGAGTTG	CAGTGAAGCGAGATCAACCACTGCATCCAGCTGGGGGAGAGAGCGA	148439
QY	2129	GACTTCCTCTCA	AAAAAACAACCAAAAGAAATTA	2165
Db	148440	AATTTGATCT	CAAAAAAACAACCAAAAGAAATTA	148476
RESULT 15				
AC005037/c				
LOCUS	AC005037	190508 bp	DNA	linear
DEFINITION	Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.			
ACCESSION	AC005037			
VERSION	AC005037.2	GI:4827310		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Sulston, J.E. and Waterston, R.			
JOURNAL	Toward a complete human genome sequence			
MEDLINE	99063792			
PUBMED	9847074			
REFERENCE	2 (bases 1 to 190508)			
AUTHORS	Abbott, A. and Le, T.			
TITLE	The sequence of Homo sapiens BAC clone RP11-469M7			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 190508)			
AUTHORS	Waterston, R.H.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-JUN-1998) Genome Sequencing Center, Washington University, School of Medicine			

RESULT 15	REFERENCE
AC005037/c	AUTHORS
LOCUS	TITLE
DEFINITION	JOURNAL
ACCESSION	MEDLINE
VERSION	PUBMED
KEYWORDS	REFERENCE
SOURCE	AUTHORS
ORGANISM	TITLE
	JOURNAL
	REFERENCE
	AUTHORS
	TITLE
	JOURNAL

REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 190508)
TITLE	Sulston, J.P. and Waterston, R.
JOURNAL	Toward a complete human genome sequence
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)
PUBMED	99063792
REFERENCE	9847074
AUTHORS	2 (bases 1 to 190508)
TITLE	Abbott, A. and Le, T.
JOURNAL	The sequence of Homo sapiens BAC clone RP11-469M7
REFERENCE	Unpublished
AUTHORS	3 (bases 1 to 190508)
TITLE	Waterston, R.H.
JOURNAL	Direct Submission
	Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis

```

MO 63108, USA
4 (bases 1 to 190508)
Waterston,R.H.
Direct Submission
Submitted (14-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 190508)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 14, 1999 this sequence version replaced gi:3309089.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0469M07
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-13J8; the clone sequenced to the right is RP11-9JMS. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES

```

source
1. 190508
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-469M7"
   /clone_lib="RPCI-11"
repeat_region
155..222
   /rpt_family="L2"
repeat_region
982..1299
   /rpt_family="Alu"
repeat_region
1502..1533
   /rpt_family="(CA)n"
repeat_region
1535..1823
   /rpt_family="Alu"
repeat_region
2005..2308
   /rpt_family="Alu"
repeat_region
2392..2562
   /rpt_family="Alu"
repeat_region
3284..3315
   /rpt_family="L1"
repeat_region
3316..3338
   /rpt_family="Alu"
repeat_region
3339..3621
   /rpt_family="(CMAA)n"
repeat_region
3650..3743
   /rpt_family="Alu"
repeat_region
4355..4684
   /rpt_family="AT-rich"
repeat_region
4727..5018
   /rpt_family="Alu"
repeat_region
5545..5585
   /rpt_family="Alu"
repeat_region
5586..5873
   /rpt_family="(CATA)n"
repeat_region
5876..6007
   /rpt_family="Alu"
repeat_region
6327..6713
   /rpt_family="Alu"
repeat_region
6924..7294
   /rpt_family="MER21_g"
repeat_region
8027..8049
   /rpt_family="MER2_type"
repeat_region
8050..8354
   /rpt_family="AT-rich"
repeat_region
8355..8661
   /rpt_family="Alu"
repeat_region
8356..8684
   /rpt_family="Alu"
repeat_region
9450..9752
   /rpt_family="AT-rich"
repeat_region
10371..10880
   /rpt_family="Alu"
repeat_region
10762..11038
   /rpt_family="Alu"
repeat_region
11039..11109
   /rpt_family="Alu"
repeat_region
11138..11410
   /rpt_family="Alu"
repeat_region
11446..11663
   /rpt_family="L1"
repeat_region
11664..11959
   /rpt_family="Alu"
repeat_region
11960..12189
   /rpt_family="L1"
repeat_region
12190..12334
   /rpt_family="Alu"
repeat_region
12335..12360
   /rpt_family="L1"
repeat_region
12464..12514
   /rpt_family="L1"
repeat_region
12515..12822
   /rpt_family="Alu"
repeat_region
12823..12862
   /rpt_family="L1"
repeat_region
12857..13096
   /rpt_family="L1"
repeat_region
13101..13330
   /rpt_family="L1"
repeat_region
13331..13634
   /rpt_family="MaLR"
repeat_region
13635..13692
   /rpt_family="Alu"
repeat_region
13704..13990
   /rpt_family="MaLR"
repeat_region
14082..14387
   /rpt_family="L1"
repeat_region
14402..14704
   /rpt_family="Alu"

```

```
QY 1981 ATACAAAAATTAGCTAGTGTGGTGGCGCATGCCCTGTAATCCCACTACTTGGGAGGCT 2040
DB |||||
QY 1981 ATACAAAAATTAGCTAGTGTGGTGGCGCATGCCCTGTAATCCCACTACTTGGGAGGCT 2040
DB |||||
QY 2041 GAGGCAGGAGATTTCTTGAACCCGGGAGGCGAGGTTGAGTGAAGCGAGATCACCA 2100
DB |||||
QY 2041 GAGGCAGGAGATTTCTTGAACCCGGGAGGCGAGGTTGAGTGAAGCGAGATCACCA 2100
DB |||||
QY 2101 CTGCATCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAAACA 2160
DB |||||
QY 2101 CTGCATCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAAACA 2160
DB |||||
QY 2161 ATTAGCAAAATTAGACATTGACAGAGAACCTGAAGGGGTCAGACCAGTACATTTT 2220
DB |||||
QY 2161 ATTAGCAAAATTAGACATTGACAGAGAACCTGAAGGGGTCAGACCAGTACATTTT 2220
DB |||||
QY 2221 TGTGCCACATGCCCAAGTACTTCTGAGGATGATGCTGGATGAGCTGCCATCTGAAATCA 2280
DB |||||
QY 2281 TCCAGTCTTGTTCAGAACTTTTCAACCCGAGCAGGAGCCAGGACTGGAAATCGAGTCTCCT 2340
DB |||||
QY 2281 TCCAGTCTTGTTCAGAACTTTTCAACCCGAGCAGGAGCCAGGACTGGAAATCGAGTCTCCT 2340
DB |||||
QY 2341 GGTCACTGGCCAGAGAGTTGGCCTTGACCTTGAGACCAGTGGCCAAACAAAGGAGCTGCTT 2400
DB |||||
QY 2341 GGTCACTGGCCAGAGAGTTGGCCTTGACCTTGAGACCAGTGGCCAAACAAAGGAGCTGCTT 2400
DB |||||
QY 2401 AGTCTACTCCAGGAATCCAGGTGCTTGTCTTCTCGGAAAGTGAATCAATTGGCCGAG 2460
DB |||||
QY 2401 AGTCTACTCCAGGAATCCAGGTGCTTGTCTTCTCGGAAAGTGAATCAATTGGCCGAG 2460
DB |||||
QY 2461 CACTCCGTATTTTCTCTCTTCCAGGGGAGGATCCTTAGGGCAGTATTTGGGAAAGACA 2520
DB |||||
QY 2461 CACTCCGTATTTTCTCTCTTCCAGGGGAGGATCCTTAGGGCAGTATTTGGGAAAGACA 2520
DB |||||
QY 2521 TGGGCTTGAAGACACCCGGGTGAATGTCATAGCTTGCCTGGTCTGAGCTCTCATGGTAA 2580
DB |||||
QY 2521 TGGGCTTGAAGACACCCGGGTGAATGTCATAGCTTGCCTGGTCTGAGCTCTCATGGTAA 2580
DB |||||
QY 2581 GGTCTCTACAGACACCGAAAGATGGGGGCA CAGGACAGATCAGTAGGGTCAGAGCATC 2640
DB |||||
QY 2581 GGTCTCTACAGACACCGAAAGATGGGGGCA CAGGACAGATCAGTAGGGTCAGAGCATC 2640
DB |||||
QY 2641 TCAGGACCCGAGGGCAATATGTCCTGACGAGGATTAAGAGCTTGGGCTCTCATATGCT 2700
DB |||||
QY 2641 TCAGGACCCGAGGGCAATATGTCCTGACGAGGATTAAGAGCTTGGGCTCTCATATGCT 2700
DB |||||
QY 2701 GTTTCGGGCTCAACTGGCAGCTCCGTCACTTACTGGTGTGCTGACCATGGGCAAGTTA 2760
DB |||||
QY 2701 GTTTCGGGCTCAACTGGCAGCTCCGTCACTTACTGGTGTGCTGACCATGGGCAAGTTA 2760
DB |||||
QY 2761 TTCCATCTCTCCATATCTTCTTCCATCTTTTAAATGGAATATGGGGTACCCACCTCC 2820
DB |||||
QY 2761 TTCCATCTCTCCATATCTTCTTCCATCTTTTAAATGGAATATGGGGTACCCACCTCC 2820
DB |||||
QY 2821 CAGGTCTACAGAGGCTTACAGAAAAAGATTTGTAATTTGGCTTCAGTAAATTTTC 2880
DB |||||
QY 2821 CAGGTCTACAGAGGCTTACAGAAAAAGATTTGTAATTTGGCTTCAGTAAATTTTC 2880
DB |||||
QY 2881 AATACCTCCAGCTATTCTTATTCACATCCAGCCCTTTCGCTGTGCTGGGTGAAAA 2940
DB |||||
QY 2881 AATACCTCCAGCTATTCTTATTCACATCCAGCCCTTTCGCTGTGCTGGGTGAAAA 2940
DB |||||
QY 2941 CACATGTCAGTGTTCCTGAGCGTTTCCAAAGAGAGATTCCTCAAAATTAACCTGCCAG 3000
DB |||||
QY 2941 CACATGTCAGTGTTCCTGAGCGTTTCCAAAGAGAGATTCCTCAAAATTAACCTGCCAG 3000
DB |||||
QY 3001 TCTGAAGAATCTCAAAACATCCCGCAGCATCTCTGAGGGCGGGCTTGGGATGGGAC 3060
DB |||||
QY 3001 TCTGAAGAATCTCAAAACATCCCGCAGCATCTCTGAGGGCGGGCTTGGGATGGGAC 3060
DB |||||
```

```
QY 3061 TCCCCGCCCGGGTCTCTGAACAGGATCGGTGCGCAGGACACACACACAGCCAGCCTG 3120
DB |||||
QY 3061 TCCCCGCCCGGGTCTCTGAACAGGATCGGTGCGCAGGACACACACACAGCCAGCCTG 3120
DB |||||
QY 3121 TGTGTGCGGGCGGAGTCCGGTTCGGGTGAGCAGCGCTGCTGTTGGGCGGGC 3180
DB |||||
QY 3121 TGTGTGCGGGCGGAGTCCGGTTCGGGTGAGCAGCGCTGCTGTTGGGCGGGC 3180
DB |||||
QY 3181 AGAGCAATTGTTTCGAGGCGTACCGAGCCCCCGCGCTCGCCCGGAGGAGCGGGCT 3240
DB |||||
QY 3181 AGAGCAATTGTTTCGAGGCGTACCGAGCCCCCGCGCTCGCCCGGAGGAGCGGGCT 3240
DB |||||
QY 3241 TCCCGCGTCCCAAGCTCCAGATCCTTGGGGTGGCTGCCACGCTCTCCCTTCCACGCGCTG 3300
DB |||||
QY 3241 TCCCGCGTCCCAAGCTCCAGATCCTTGGGGTGGCTGCCACGCTCTCCCTTCCACGCGCTG 3300
DB |||||
QY 3301 GGGGGAAGGGAAGACGGGACGGAGATGTTAGTGTGGGCGCCCGAGGGTTTCACTACT 3360
DB |||||
QY 3301 GGGGGAAGGGAAGACGGGACGGAGATGTTAGTGTGGGCGCCCGCGGGTTTCACTACT 3360
DB |||||
QY 3361 GTTTCCTGAGAACTTCCCGAGTCCCGAGCCACCCCGCTTCTCCGTGTCGAGGGCGGT 3420
DB |||||
QY 3361 GTTTCCTGAGAACTTCCCGAGTCCCGAGCCACCCCGCTTCTCCGTGTCGAGGGCGGT 3420
DB |||||
QY 3421 CTTGGGCTAGGCTCCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3480
DB |||||
QY 3421 CTTGGGCTAGGCTCCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3480
DB |||||
QY 3481 GCTTCCCGACCGGGATGCGGGGACAGGCGCCAGCGGGTGAAGAGAGCTTGAAGG 3540
DB |||||
QY 3481 GCTTCCCGACCGGGATGCGGGGACAGGCGCCAGGCGGGTGAAGAGAGCTTGAAGG 3540
DB |||||
QY 3541 AGAAACAGAGGGGAGGGGAGCGAGAGCTGCGGCGAGAGGAAACAGCGAATTCGCGCGA 3600
DB |||||
QY 3541 AGAAACAGAGGGGAGGGGAGCGAGAGCTGCGGCGAGAGGAAACAGCGAATTCGCGCGA 3600
DB |||||
QY 3601 GCCAATGGCAACCGGACGAGGAGTGGCACCAGTTCCTTCCGCGCAATGACGAGCGCGA 3660
DB |||||
QY 3601 GCCAATGGCAACCGGACGAGGAGTGGCACCAGTTCCTTCCGCGCAATGACGAGCGCGA 3660
DB |||||
QY 3661 GTTTACAGAGCTCATTTAGCAATTTCCCGAGGCGAGGGGACAGGCGCGGGTGG 3720
DB |||||
QY 3661 GTTTACAGAGCTCATTTAGCAATTTCCCGAGGCGAGGGGACAGGCGCGGGTGG 3720
DB |||||
QY 3721 TGTGTGTGCGGTGTCGCGAGCATCCCGCGCGCTTGTGCGGTTCGCGGAGCTTCGGCC 3780
DB |||||
QY 3721 TGTGTGTGCGGTGTCGCGAGCATCCCGCGCGCTTGTGCGGTTCGCGGAGCTTCGGCC 3780
DB |||||
QY 3781 TGTGTGTCTCTCCCTTCCCGCTTACCTCCAGCGGACCGCGCGCTTCACTTCC 3840
DB |||||
QY 3781 TGTGTGTCTCTCCCTTCCCGCTTACCTCCAGCGGACCGCGCGCTTCACTTCC 3840
DB |||||
QY 3841 TCGCACTTTGCGCTTGTGCGAGCGGATTAAGAGGGGCTGAGGAAATACCGGACAGGT 3900
DB |||||
QY 3841 TCGCACTTTGCGCTTGTGCGAGCGGATTAAGAGGGGCTGAGGAAATACCGGACAGGT 3900
DB |||||
QY 3901 CACCCGTTCG 3910
DB |||||
QY 3901 CACCCGTTCG 3910
DB |||||
```

RESULT 2

ABK94819

ID ABK94819 standard; DNA; 1970 BP.

XX AC ABK94819;

XX AC

XX AC

DT 29-AUG-2002 (first entry)

XX Human stearyl-CoA desaturase (SCD) control region.

XX DE Fatty acid regulated gene; polyunsaturated fatty acid disorder;

XX DE PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;

KW

KW

KW	dyslipidaemia; atherosclerosis; coronary artery disease; cerebrovascular disease; peripheral vascular disease; inflammation; sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis; acne; body weight disorder; obesity; cachexia; anorexia; psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome; diabetes; diabetic complication; genetic polymorphism; ds.
OS	Homo sapiens.
XX	WO200240666-A2.
PN	23-MAY-2002.
XX	19-NOV-2001; 2001WO-CA001632.
PF	17-NOV-2000; 2000US-0248589P.
XX	(XENO-) XENON GENETICS INC.
PA	Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ; Ponton A, De Antueno RJ, Jenkins DK, Nwaka SO; WPI; 2002-508327/54.
DR	Novel isolated polypeptide segment encoded by fat regulated genes, useful for diagnosing the presence of or a predisposition for a disorder involving fatty acid regulated genes in a subject.
PT	Claim 23; Fig 26; 225pp; English.
PT	The invention describes an isolated polypeptide segment (I) whose genes are fat regulated. (I) or the polynucleotide encoding it (II) are useful for diagnosing the presence of or a predisposition for a disorder involving fatty acid regulated genes in a subject. A composition containing (I) or (II) is useful for treating a disorder involving fatty acid regulated genes, where the disorder is selected from a polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis, coronary artery disease, cerebrovascular disease or peripheral vascular disease), inflammation (such as sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis or acne), body weight disorders (such as obesity, cachexia or anorexia), psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic complications. (I) or (II) is useful as research agent and materials for discovery of treatments and diagnostics for a disease, particularly human disease. (II) is useful for constructing nucleotide probes and primers, for detecting genetic polymorphism, for detecting changes in the level of expression of (II), and as a diagnostic tool. This sequence represents a control region from a gene encoding a fatty acid regulated protein
CC	Sequence 1970 BP; 428 A; 590 C; 588 G; 364 T; 0 U; 0 Other;
XX	Query Match 44.8%; Score 1751.8; DB 6; Length 1970; Best Local Similarity 99.7%; Pred. No. 0; Matches 1786; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY	2121 GAGAGCGAGACTTCCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 2180
Db	1 GAGAGCGAGACTTCCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 60
QY	2181 CAGAGAACTGGAAGGGGTGCACACGTACAGATTTCTGTGCCATGCCAAGTACT 2240
Db	61 CAGAGAACTGGAAGGGGTGCACACGTACAGATTTCTGTGCCATGCCAAGTACT 120
QY	2241 TCTGAGCATGACTGGATGAGCTCTCCACATCTCAAAATCATCCAGTCTTGTTCAGACTT 2300
Db	121 TCTGAGCATGACTGGATGAGCTCTCCACATCTCAAAATCATCCAGTCTTGTTCAGACTT 180
QY	2301 TCACACGGACACGGAGCCAGGACTGGAATGCGAGTCTCTCGTGCACGACAGAGTTG 2360
Db	181 TCACACGGACACGGAGCCAGGACTGGAATGCGAGTCTCTCGTGCACGACAGAGTTG 240
QY	2361 GCCTTGACCTTGAGACGAGTGGCCACAAGGAGTGCTTTAGTCTACCTCCCAAGAAATC 2420

Db 64 GGTGAAGACAGCTGAGAGGAGAAACAGAGGGGAGGAGCGAGCTGGCGGAG 123
QY 3579 AGGGAACAGAGATTGGCCGAGCCAAATGCAACGGCAGGACGAGTGCGACCAAAATTC 3638
Db 124 AGGGAACAGAGATTGGCCGAGCCAAATGCAACGGCAGGACGAGTGCGACCAAAATTC 183
QY 3639 CTTCCGGCAATACGAGCGGAGTTCACAGAGCTTATAGCATTTCCCGAGGCGAGG 3698
Db 184 CTTCCGGCAATACGAGCGGAGTTCACAGAGCTTATAGCATTTCCCGAGGCGAGG 243
QY 3699 GCGAGGGCAGAGCGCGGGTGTGTGTTGTCGGCAGCATCCCGCGCCCTGCT 3758
Db 244 GCGAGGGCAGAGCGCGGGTGTGTGTTGTCGGCAGCATCCCGCGCCCTGCT 303
QY 3759 GCGGTGCGCGAGCTCGGCTCTGCTCTCCCTCCGCGCTTACCTCACGCGGG 3818
Db 304 GCGGTGCGCGAGCTCGGCTCTGCTCTCCCTCCGCGCTTACCTCACGCGGG 363
QY 3819 ACCGCCCGCCAGTCACTCTCGCACTTGTGCTTGGCAGCGATAAAGGGGG 3878
Db 364 ACCGCCCGCCAGTCACTCTCGCACTTGTGCTTGGCAGCGATAAAGGGGG 423
QY 3879 CTGAGGAATACCGGACAGCTCACCGTTCGCC 3910
Db 424 CTGAGGAATACCGGACAGCTCACCGTTCGCC 455

RESULT 4
AAK83253
ID AAK83253 standard; DNA; 8880 BP.
XX AC
XX AAK83253;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38065.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246527P.

PR	08-NOV-2000;	2000US-0246528P.	14	TTTAAATCATACAAATTTAATGCTTTTATGCTTTTATGATATATTCACAG-GTTGTGATCCAT---CAC	69
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.	6137	TTTAAAGTGTGCAATTTGATTAATTTTGTGATTTTCTCAGAGTTTGTGAACCAATTTACCAC	6196
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
PR	17-NOV-2000;	2000US-0249207P.	70	AATCCATTTTGAACAGATTTTATTACTCCAAAATAAAACCCCTGTCATTCCTTAGCCATCAC	129
PR	17-NOV-2000;	2000US-0249208P.			
PR	17-NOV-2000;	2000US-0249209P.	6197	CATCAAGTTTGAACAATTTTATAACCCAGAAAGAACCCCTGTAGCCCTTTAGCTGTAC	6256
PR	17-NOV-2000;	2000US-0249210P.			
PR	17-NOV-2000;	2000US-0249211P.	130	CCCCCA--ACATCTCCATCTCTTCCAAAGCCCTGGGCAACCAACCAATCTACTTTCTGT	187
PR	17-NOV-2000;	2000US-0249212P.			
PR	17-NOV-2000;	2000US-0249213P.	6257	CTTCAAGTACCACTTTCTCTCTCCCTTTCCCTTAGCTATCATATAAACTTTCTGT	6316
PR	17-NOV-2000;	2000US-0249214P.			
PR	17-NOV-2000;	2000US-0249215P.	188	CTCTATAAATTTGCCAATTTCTGGACATTTCTATAAATGGAAGCAACAACATGTGAGAC	247
PR	17-NOV-2000;	2000US-0249216P.			
PR	17-NOV-2000;	2000US-0249217P.	6317	CTATATAGATTTGGCTATCTCTGGACATTTCTATAAATGGAATCATATAATGTGGCT	6376
PR	17-NOV-2000;	2000US-0249218P.			
PR	17-NOV-2000;	2000US-0249244P.	248	TTTGTGACTGGCTGCTTTCTACTTAGCAATTTATTTTAAAGGCTCA-TTATGTTACAGTAC	306
PR	17-NOV-2000;	2000US-0249245P.			
PR	17-NOV-2000;	2000US-0249246P.	6377	TTTCTGACTAGCTTCTTTTCTACTTAGCGTGATCATTTCAAGGTTTCATTTATGTTAGCAT	6436
PR	17-NOV-2000;	2000US-0249265P.			
PR	17-NOV-2000;	2000US-0249297P.	307	TTAGCAGTACTTCAATCTTTTATTTTCTCAAAATGATTTTCACTGTGTGGGTATCCCATTA	366
PR	17-NOV-2000;	2000US-0249299P.			
PR	17-NOV-2000;	2000US-0249300P.	6437	GTATCAGTGTGTTTGTGTTTGTGTTGTCAGACGGAGTTTGTCTCT-----	6482
PR	01-DEC-2000;	2000US-0250160P.			
PR	01-DEC-2000;	2000US-0250391P.	367	TCATATTATTAGACAGAGTTTCTCTCTCTCACCAGGCTGGAGTGCAGTGGCAGCAATC	426
PR	05-DEC-2000;	2000US-0251030P.			
PR	05-DEC-2000;	2000US-0251988P.	6483	-----TCTTGGCCAGGCTGGAGTGCAGTGGCGAGTC	6514
PR	06-DEC-2000;	2000US-0251479P.			
PR	08-DEC-2000;	2000US-0251856P.			
PR	08-DEC-2000;	2000US-0251868P.			
PR	08-DEC-2000;	2000US-0251869P.			
PR	08-DEC-2000;	2000US-0251989P.			
PR	08-DEC-2000;	2000US-0251990P.			
PR	11-DEC-2000;	2000US-0254097P.			
PR	05-JAN-2001;	2001US-0259678P.			
XX					
PA	(HUMA-) HUMAN GENOME SCI INC.				
XX					
PI	Rosen CA, Barash SC, Ruben SM;				
XX					
XX	WPI; 2001-483426/52.				
DR					
XX					
XX					
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,				
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.				
XX					
XX	Disclosure; SEQ ID NO 38065; 307lpp + Sequence Listing; English.				
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)				
CC	amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic				
CC	activity, and can be used in gene therapy and vaccine production. (I)				
CC	proteins and polynucleotides may be used in the prevention, diagnosis and				
CC	treatment of diseases associated with inappropriate (I) expression. For				
CC	example, they may be used to treat disorders associated with decreased				
CC	expression by rectifying mutations or deletions in a patient's genome				
CC	that affect the activity of (I) by expressing inactive proteins or to				
CC	supplement the patients own production of (I). Additionally, (I)				
CC	polynucleotides may be used to produce the secreted (I), by inserting the				
CC	nucleic acids into a host cell and culturing the cell to express the				
CC	protein. (I) proteins and polynucleotides may be used to prevent,				
CC	diagnose and treat immune/haematopoietic-related diseases, especially				
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703				
CC	to AAK67694 represent human immune/haematopoietic antigen genomic				
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK62169				
CC	represent sequences used in the exemplification of the present invention				
XX					
XX	Sequence 8880 BP; 2388 A; 1807 C; 1844 G; 2841 T; 0 U; 0 Other;				
SQ					

Query Match 9.9%; Score 386; DB 4; Length 8880;
Best Local Similarity 54.7%; Pred. No. 2.5e-69;
Matches 1194; Conservative 0; Mismatches 850; Indels 138; Gaps 16;

1064 CCACCTTTTGGACCATTAATGAATTAATCTCCAGTGAATTAATCATGT---ATACATTTGTGT 1120
 7169 CTACCTTTTGGATTAATGAATTAATCTCTGGAACATTCGTATACAGATTTTAATGT 7228
 1121 GGGCATATGTTTTCATCTCTGTTGGTATTAATCTAGGAGTGAATTTGCTGATCCGGG 1180
 7229 GGATCTATGTTTTCATCTCTGTTGGTATTAATCTAGGAGTGAATTTGCTGATG 7288
 1181 TAATATTTTTCAGAGGAGTTCAGGGGAGAGAAATCTGGGAAATGAAGCATGTTAG 1240
 7289 GTGACTCTGTAGAGCTTCAAGCTCCGTAATGTTTCCAGAGTGGCTGCACATTTTCC 7348
 1241 AAATCAGCAAGAGTGCAGGGGTTTTCGAGATTTTATTTATTTATTTCTGTTGACAAATGG 1300
 7349 ATTGATGTTGAACATCTTTTCAGGCTTATCGACATTTGATATCTTTGGAGAAATGC 7408
 1301 CAGTTTGTATGAAGATCAAGTTTATTAAGTGAGAGTGAATTAAGGCTGGAATPAGG 1360
 7409 CCAATCGGTTTCTTCTGCTGATTTTAAATTAATTTATTTTATTTACTGAGTTGTAAGC 7468
 1361 CGTTCAGAGTAAATCAATCAAGCACTTTCGAATACCAAAATTAAGGAGCTGGCTGTAAAC 1420
 7469 AATTTCTTTAGATAGTCCAGATCAAGTCTTTATCAGA-----TATAT 7511
 1421 AAAATAATAAAAAATCACAAATTTTTCGAGATTTTTCGAGAAAGAGTCTTCTTTTCC 1480
 7512 TATTTGCAATATTTTCTCCCATTTTGTGGTGTCTTCGATTTCTTTGCTTTTTCOA 7571
 1481 CTGGCTGGAGGAGTGGTGTGATCTCAGCTCAGTCAACTTTCCGCTCCCGGTTCAAG 1540
 7572 TTG-----GAGTAAATATATACATAACATAAAATTAACATTTTGGGGTGGGGCAAT 7625
 1541 CAATTTCTCTGCTTCAGGCTCCCAAGTAGCTGGAGTACAGGACTTCCACCATGCCA 1600
 7626 GGTTCACACTGTAAATCCAGCAATTTAGAGGAGGAGGTAGGAGATGCTTCAGCCCA 7685
 1601 GCTGATTTTGTATTTTATAGTAGAGTGGATTTCACTTTGTTGGCCAGCTGTCTCAA 1660
 7686 GGAGTTTGAGACCACTCTGGGCAACATAGCAAGATCTGTCTACAA----- 7734
 1661 ACTTTTCTGTCAATAATTTGTGTAATATGTTCTTTGTTGAGGTAGGGCCCCCAGA 1720
 7735 -----ATATTAATAAATTAGCTGGGCTGGCTGCACACTTTCTAGTCCAGGA 7781
 1721 CCAAAAAAATAAATCTTAGAATCAAAATCACTGTGTTGTTGACCACTGTCACTTCA 1780
 7782 TGAGGTGAGAAATCGCTTGAGTCAGAGGTGAGGCTGTAAAGTACCGGTGATCAC 7841
 1781 AACCAAGTGTGACAGGGCTCAGGAGTAGAGTGTCTCTGCTCGAAGAGAGAAATAGA 1840
 7842 CACTGCAGTCCAGGCTG-----ATGGCAGAGTGAGACTCTGTATTTTAAAAA 7894
 1841 ATGAAAAATATTTCTCGGGCCAGGGTGTGCTCATGCTGTATCTTATCCAGCACTTTGGA 1900
 7895 AACAAAAAACAAGAGCCAGGGGGTGGCTCAAGCTGTATTCACAGCACTTTGGG 7954
 1901 GGCCAGGATGTGGATCACCTGAGGTGAGGATTTCAAAACAGGCTGGCCAAACATGTG 1960
 7955 GACCAAGGGGCGAGATCACTGAGTGGAGTTCGAGACAGGCTGACCTACACGAG 8014
 1961 AAACCCGCTCTCTACTAAAAATAAAAAATTAAGTGTGGTGGCGCATGCTGTAA 2020
 8015 AAACCGTCTCAACTGAAAAATTC-----AAATTTGCGAGTGTGGTGGCGCATGCTGTAA 8072
 2021 TCCAGCTACTTTGGAGGCTGAGCGAGGAGATTTCTTGAACCCGGAGGAGGAGGTTC 2080
 8073 TCCAGCTATTTCAGAGGTTGAGGAGAGAGATCGTTTGAACCTGGGAGGAGGTTTC 8132
 2081 AGTGAAGCGAGATCACACCACTGCACTCCAGCTGGGGGA--GAGAGGAGGACTTCTCTC 2139
 8133 GGTGAGCCAGATCGACCATTTGACCTTAGCTTGGGCAACAGAGGCAACTCCGCTC 8192
 2140 AAAAAACAAAAACAAAAAGAA 2161

Db 8193 AAAAAAAAAAAAAAGAGAAA 8214

RESULT 5

ABQ8096/c

ID ABQ8096 standard; cDNA; 92638 BP.

XX ABQ8096;

XX AC

XX DT 18-SEP-2002 (first entry)

XX Human osteoblast differentiation related cDNA SEQ ID NO 3.

DE Human; osteoblast; stem cell differentiation; bone tissue deposition;

XX osteoporosis; osteopathic; ss.

XX Homo sapiens.

XX WO200250301-A2.

XX 27-JUN-2002.

XX 18-DEC-2001; 2001WO-US048276.

XX 18-DEC-2000; 2000US-0255882P.

XX 24-APR-2001; 2001US-0285691P.

XX (GENE-) GENE LOGIC INC.

XX (PROC) PROCTER & GAMBLE CO.

XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

XX Mertz L;

XX WPI; 2002-557663/59.

Use of genes and their expression profiles associated with osteoblast

differentiation for screening modulators bone formation, for diagnosing

or treating e.g. osteoporosis, or as markers for the differentiation

process.

Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.

The invention relates to genes and their expression profiles are used

for: (a) screening modulators of precursor stem cell differentiation into

osteoblasts, or bone tissue deposition; (b) diagnosing abnormal

deposition of bone tissue, abnormal rate of osteoblast formation or

osteoporosis; or (c) treating or monitoring treatment of the conditions

cited in (b), or monitoring the progression of bone tissue deposition.

Specific conditions include postmenopausal osteoporosis, glucocorticoid

osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-

induced abnormalities in bone formation or bone loss, conditions that

involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),

skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome

or fibrous dysplasia. The present sequence is that of an osteoblast

differentiation associated cDNA marker of the invention. Note: The

sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences

Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;

Query Match

Best Local Similarity 9.6%; Score 375.4; DB 6; Length 92638;

Matches 1046; Conservative 0; Mismatches 756; Indels 115; Gaps 12;

Qy 366 ATCATATATTAGAGACAGGTTCTCTCTGTCAACCAGGCTGGAGTGCATCAAT 425

Db 57931 AGCAATTTTTTGGACAGGTTCTCTCTGTCACTCAGGCTGGAGTGCATCCCA 57872

Qy 426 CATAGCTCACTGTAACTCAAACTCTCGGGCTCAAGTGATCCCTACTACCTCAGCTCCAG 485

Db 57871 CATAGCTCACTGGCAGCTCGACCTTCAGGGCTCAAGCAATCTCTCGCTCAACCTCCA 57812

us-10-016-725-15_copy_1_3910.rng

Fri May 14 12:58:33 2004

PR	02-MAR-2000;	2000US-0186350P.	PR	02-OCT-2000;	2000US-0237040P.
PR	16-MAR-2000;	2000US-0189874P.	PR	13-OCT-2000;	2000US-0239935P.
PR	17-MAR-2000;	2000US-0190076P.	PR	13-OCT-2000;	2000US-0239937P.
PR	18-MAR-2000;	2000US-0190123P.	PR	20-OCT-2000;	2000US-0240960P.
PR	19-MAY-2000;	2000US-0205513P.	PR	20-OCT-2000;	2000US-0241221P.
PR	07-JUN-2000;	2000US-0209467P.	PR	20-OCT-2000;	2000US-0241785P.
PR	28-JUN-2000;	2000US-0214886P.	PR	20-OCT-2000;	2000US-0241787P.
PR	30-JUN-2000;	2000US-0215135P.	PR	20-OCT-2000;	2000US-0241808P.
PR	07-JUL-2000;	2000US-0216647P.	PR	20-OCT-2000;	2000US-0241809P.
PR	07-JUL-2000;	2000US-0216880P.	PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000;	2000US-0217487P.	PR	01-NOV-2000;	2000US-0244617P.
PR	11-JUL-2000;	2000US-0217496P.	PR	08-NOV-2000;	2000US-0246474P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246476P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR	18-AUG-2000;	2000US-0226279P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-02323063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	26-SEP-2000;	2000US-0235484P.	PA	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235834P.	PI	WPI; 2001-483426/52.	
PR	27-SEP-2000;	2000US-0235836P.	XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides.	
PR	29-SEP-2000;	2000US-0236327P.	XX	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	29-SEP-2000;	2000US-0236367P.	DR	Disclosure; SEQ ID NO 23856; 3071pp + Sequence Listing; English.	
PR	29-SEP-2000;	2000US-0236368P.	XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	2000US-0236370P.	XX		
PR	02-OCT-2000;	2000US-0236802P.	XX		
PR	02-OCT-2000;	2000US-0237037P.	XX		
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	CC		

CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 5491 BP; 1662 A; 1226 C; 1439 G; 1164 T; 0 U; 0 Other;

Query Match 9.4%; Score 368.8; DB 4; Length 5491;
Best Local Similarity 56.0%; Pred. No. 7.5e-66;
Matches 1023; Conservative 0; Mismatches 627; Indels 178; Gaps 10;

QY 361 CCCATATCATATTATTAGACAGGTTCTCACTCTGTCCACCCAGGCTGGAGTGCAGTGGC 420
DB |||||
QY 3658 CCCTAATTAATTTTTTGAGACGGTGTCTCACTCTGTGCCCCAAGTTGGAGTGAATGAC 3599
DB |||||
QY 421 ACAATCATAGTCTACTGTAACCTCAAACTCTGGGCTCAAGTATCTTACCTCAGCC 480
DB |||||
QY 3598 TGAATCTTGGCTCACTGCAACCTCCACCTAGCGGGTCAAGTGAATCTCCCGCCTCAGCC 3539
DB |||||
QY 481 TCCAGAGTAGTACTAGTACAGGACACACAGCCATCTCGCTAAATTTTTTTTAAAT 540
DB |||||
QY 3538 TCCCAAGTAGTGGGACTATTAGGCGCGCACCATCAGCTTGGCTTAATTTTGTAT 3483
DB |||||
QY 541 TTTTCATTTTATGTAATTCATTTCTTTCTTTTGTGTTGTTGTTGAGATAGGCTC 600
DB |||||
QY 3482 -----TTTAGTAAAGACAAGTTT 3463
DB |||||
QY 601 ACTTTGTTTACCAGGCTGGAGGCGAGTGGCATGTTGAGCTGAGCAGCCTTGACTTCT 660
DB |||||
QY 3462 GCCATGTTGGCCAGCTGGT -----CTTGAATCAT 3432
DB |||||
QY 661 GGGCTCAAGTATCTCTGCTCCTCAGCTCCCAAGTAGTGGACTACAAAACACGTGCA 720
DB |||||
QY 3431 GACCTCAGGTGATCGGCTCACATCAGCTCCCAAGTGTGGGATACAGGCGTAGCCA 3372
DB |||||
QY 721 CCATGCTGGCTGATA -----TTTTTTTCTTGAACAGGGTATCA 761
DB |||||
QY 3371 CTGTGCTCTGCCATTAATTTATTTATTTATTAATTTTATTTTATTTTATAGACAGTCTCA 3312
DB |||||
QY 762 CTCTGTGCTCCAGGCTGGAGTACAGTGGGCTTAATTAAGTCTCACTCAGCAGCTCC-CTTCC 820
DB |||||
QY 3311 CTCTGTGCTCCAGGCTGGAGTGGAGCAATCTCGGCTCAGTGAACCTCCATCTCC 3252
DB |||||
QY 821 TGGGCTCAGCAATCGCTGGCTCAGCATCTGTAGTACTGGGACTCAGGCTTGTGCC 880
DB |||||
QY 3251 TAGGTTCAAGCTATTCTCTGCTTACGCTCCCAAGTATTTGGATTAAGGGGCCACC 3192
DB |||||
QY 881 ACCAGGCCAGCTAAGTTTAAATAATGATTTTTTGGTATAGAGGAGGCTTCTGTAFTGTG 940
DB |||||
QY 3191 ACCACACCGGCTAATTTTGT -----ATTTTATAGAGATGAGGTTTCAACATGTTG 3138
DB |||||
QY 941 CTCAGGCTGTATTTTATTTGTTGAGACAGGCTCTCATATGTTGCCATGATCCCCCACC 1000
DB |||||
QY 3137 GCCAGGGGGTCAACCAACTCTGGCCTCA -----ATGATCCACTGCC 3094
DB |||||
QY 1001 TCCACTTCCCAAGTGTCTCATCTTATCTTCTTATTTAGTACAGTATTTAGGTTG 1060
DB |||||
QY 3093 TCAGCCTCCCAAGTGTCTGGATTTACAGGCGTGATCCACGCGCCGCTGATGTTTC 3034
DB |||||
QY 1061 TTTCCACTTTTACCATTATGAATAATACTCCAGTGAATTAATTCATGTAATATTTGTG 1120
DB |||||

DB 3033 TTTCTTTCTTTTATTTTCTTTTATTTTCTTTTATAGACATAGTCTTGTCTCTGTACCATGCT 2974
QY 1121 GGGCATATCTTTTTCATTTCTTGTGGTTTATATCTAGGAGTGGAAATGCTGGATCCGGG 1180
DB 2973 GAGTACGGTGGCGAGATCACGACTCAGTACAACTTTGAATTCCTGGAATCAAGCAATCC 2914
QY 1181 TAATATTTTGAAGGAGAGTTCAGGGGAAGAAACACTTGGGAAATGAAGCATGTTTAG 1240
DB 2913 TCCCATCTCAGCTCCTGAGT -----AGCTAAG 2986
QY 1241 AATCAGCAGAGTGCAGGSGTTTTTCGAGTTTTTATTTATATTTCTGTTGACAAATGTG 1300
DB 2885 GACTACAGGTGGGTACCAACCTCTGGCTGTTTTCTGAAATATCTGTTTTTAACTCC 2826
QY 1301 CAGTTTGAAGAAATCAAGTTATATCTAAAGTGAAGTGAAGTAAAGGTGGAAATAGG 1360
DB 2825 CTATTTGTTTCAATACCTCTCTCAAGAATACTATTATAGTCTCTGCTCTCCGATTT 2766
QY 1361 CGTTCAGAGTAAATCATAGACACTTTGAAATACCAAAATTAAGGAGCTTGGCTGTAAC 1420
DB 2765 TCTTGTGATCTTGCACACTGAAATTTCTGTATCTCTGGGACTGAAATTTCTT 2712
QY 1421 AAAATAATAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGAGTCTTGTCTTTCAAC 1480
DB 2711 -----TTTCTTTTTTTTTTTTTTTTTTTTTCAGACAGATACTTGTCTGTCAAC 2665
QY 1481 CTGCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTTGGCTTCCGGTTCAG 1540
DB 2664 CAGGCTAGAGTTCAGTGGCTTGATCTCAGCTCATTTGCAACCTCCGCTCTCTGGTTCAG 2605
QY 1541 CAATCTCTCTGCTTACGCTCCAGTAGCTGGAGTACAGGCACTTCCCAACCATGCCCCA 1600
DB 2604 CAATCTTGGCCTCAGCTCCCAAGTAGCTGGAAATACAGATGATGCCACACGCCCC 2545
QY 1601 GCTGATTTTTGTATTTTTTAGTAGAGATGGGATTTTCACTTTTGTGGCCAAAGTGTCTCAA 1660
DB 2544 ACTAATTTTTGTATTTTTTAGTAGGACTGGGTTTCACTGTTGGCCAGGCTGTCTGT 2485
QY 1661 ACTTTTTTGTCTGTAATTTGT -----TGAAACTATTTGTTCTTTTCTGCTGAGGTAG 1710
DB 2484 ACTCTGACCTCAAGTGTATCTGCCCTCTCCCAAGTGTCTGGGATTAAGGAT 2425
QY 1711 GGGCCCGACAGCAAAATAAATCTTAGAATCCAAATCAGTGTGTGTTTGACCACT 1770
DB 2424 GAGCCACCATGCCAGCCATGAGACTGAAATTTCTGTTGTTTTTTTGGGGTCTATTT 2365
QY 1771 GTCACTTGAGAACCAACAGTGTGACAGGCTCTCAGGAGTACAGGTGATCTCTGTCTGAAA 1830
DB 2364 TTATGTTAGGACTGTTTTCTCCAAGCTTGTGTGATCTTTGGCTCTATATTTCTCTACA 2305
QY 1831 GAGAAATGAAATGAAAATATTTCTCGGGCCAGGCTGTGCTCATGCTGTAATCCCGAG 1890
DB 2304 AGCAGGCACTAAAAGCTGGTTGAAGGCCAGACACAGCGCTCAGGCTGTATATCCAG 2245
QY 1891 CACTTTGGAGGCCCAAGCATGTGATCACTGAGGTGAGGAGTTCAAAACACAGCTGGC 1950
DB 2244 CACTTTGGAGGCTGAGGAGGAGATCA -----TGAGTTCAGGAGGTCGAGACCATCTGAC 2187
QY 1951 CAACATGTGAAACCCCGTCTCTACTAAAATAFACAAAATTAAGCTAAGTGTGTGGGCGC 2010
DB 2186 TAACTCTGGTGCACCTCTCTCTACTAAAATAFACAAAATTAAGCAGGTGTGTGGGCGG 2127
QY 2011 ATGCTGTAAATCCAGCTACTTTGGAGGTTGAGGAGGAGGAAATTTCTTGAACCCCGGAGG 2070
DB 2126 GTGCTGTAGTCCAGCTACTTTGGAGGCTGAGGAGGAGGAAATGGTATGAGGCCAGGAGG 2067
QY 2071 CAGAGGTTGCACTGAGGAGGAGATCACACCTGCTCACTCCAGCTGGGGGAGAGAGGAGA 2130
DB 2066 CGGAGCTTCAGTGTAGCCGAGATTTGCACTTGCATCTCAGCTCCAGCTGAGCAGAGCAAGA 2007
QY 2131 CTTCTCTCAAAAAAACAAAAACAAAA 2158
DB 2006 CTTCTGCTCAAAAAAACAAAAAACAAAA 1979

us-10-016-725-15 copy 1 3910.rng

Accession	Gene	Protein	Function	Ref
08-SEP-2000; 2000US-0231244P	PR	08-SEP-2000; 2000US-0231244P		
08-SEP-2000; 2000US-0231413P	PR	08-SEP-2000; 2000US-0231413P		
08-SEP-2000; 2000US-0231414P	PR	08-SEP-2000; 2000US-0231414P		
08-SEP-2000; 2000US-0232080P	PR	08-SEP-2000; 2000US-0232080P		
08-SEP-2000; 2000US-0232081P	PR	08-SEP-2000; 2000US-0232081P		
08-SEP-2000; 2000US-0231968P	PR	08-SEP-2000; 2000US-0231968P		
12-SEP-2000; 2000US-0232397P	PR	12-SEP-2000; 2000US-0232397P		
14-SEP-2000; 2000US-0232398P	PR	14-SEP-2000; 2000US-0232398P		
14-SEP-2000; 2000US-0232399P	PR	14-SEP-2000; 2000US-0232399P		
14-SEP-2000; 2000US-0232400P	PR	14-SEP-2000; 2000US-0232400P		
14-SEP-2000; 2000US-0232401P	PR	14-SEP-2000; 2000US-0232401P		
14-SEP-2000; 2000US-0233063P	PR	14-SEP-2000; 2000US-0233063P		
14-SEP-2000; 2000US-0233064P	PR	14-SEP-2000; 2000US-0233064P		
14-SEP-2000; 2000US-0233065P	PR	14-SEP-2000; 2000US-0233065P		
21-SEP-2000; 2000US-0234223P	PR	21-SEP-2000; 2000US-0234223P		
21-SEP-2000; 2000US-0234224P	PR	21-SEP-2000; 2000US-0234224P		
25-SEP-2000; 2000US-0234977P	PR	25-SEP-2000; 2000US-0234977P		
25-SEP-2000; 2000US-0234978P	PR	25-SEP-2000; 2000US-0234978P		
26-SEP-2000; 2000US-0235484P	PR	26-SEP-2000; 2000US-0235484P		
27-SEP-2000; 2000US-0235834P	PR	27-SEP-2000; 2000US-0235834P		
27-SEP-2000; 2000US-0235836P	PR	27-SEP-2000; 2000US-0235836P		
29-SEP-2000; 2000US-0236327P	PR	29-SEP-2000; 2000US-0236327P		
29-SEP-2000; 2000US-0236367P	PR	29-SEP-2000; 2000US-0236367P		
29-SEP-2000; 2000US-0236368P	PR	29-SEP-2000; 2000US-0236368P		
29-SEP-2000; 2000US-0236369P	PR	29-SEP-2000; 2000US-0236369P		
29-SEP-2000; 2000US-0236370P	PR	29-SEP-2000; 2000US-0236370P		
02-OCT-2000; 2000US-0236802P	PR	02-OCT-2000; 2000US-0236802P		
02-OCT-2000; 2000US-0237037P	PR	02-OCT-2000; 2000US-0237037P		
02-OCT-2000; 2000US-0237038P	PR	02-OCT-2000; 2000US-0237038P		
02-OCT-2000; 2000US-0237039P	PR	02-OCT-2000; 2000US-0237039P		
02-OCT-2000; 2000US-0237040P	PR	02-OCT-2000; 2000US-0237040P		
13-OCT-2000; 2000US-0239935P	PR	13-OCT-2000; 2000US-0239935P		
13-OCT-2000; 2000US-0239937P	PR	13-OCT-2000; 2000US-0239937P		
20-OCT-2000; 2000US-0240960P	PR	20-OCT-2000; 2000US-0240960P		
20-OCT-2000; 2000US-0241221P	PR	20-OCT-2000; 2000US-0241221P		
20-OCT-2000; 2000US-0241785P	PR	20-OCT-2000; 2000US-0241785P		
20-OCT-2000; 2000US-0241786P	PR	20-OCT-2000; 2000US-0241786P		
20-OCT-2000; 2000US-0241787P	PR	20-OCT-2000; 2000US-0241787P		
20-OCT-2000; 2000US-0241808P	PR	20-OCT-2000; 2000US-0241808P		
20-OCT-2000; 2000US-0241809P	PR	20-OCT-2000; 2000US-0241809P		
20-OCT-2000; 2000US-0241826P	PR	20-OCT-2000; 2000US-0241826P		
01-NOV-2000; 2000US-0244617P	PR	01-NOV-2000; 2000US-0244617P		
08-NOV-2000; 2000US-0246474P	PR	08-NOV-2000; 2000US-0246474P		
08-NOV-2000; 2000US-0246475P	PR	08-NOV-2000; 2000US-0246475P		
08-NOV-2000; 2000US-0246476P	PR	08-NOV-2000; 2000US-0246476P		
08-NOV-2000; 2000US-0246477P	PR	08-NOV-2000; 2000US-0246477P		
08-NOV-2000; 2000US-0246478P	PR	08-NOV-2000; 2000US-0246478P		

PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251088P.
 PR 05-DEC-2000; 2000US-0251719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 XX WPI; 2001-465566/50.
 XX
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 PT treating neural, immune system, muscular, reproductive, pulmonary,
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX
 PS Disclosure; SEQ ID NO 2141; 1180pp; English.
 XX
 CC The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders, including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 CC infectious disorders (e.g. Influenza). The polynucleotides of the
 CC invention can also be used in gene therapy. AAS41685-AAS42192 represent
 CC DNA sequences encoding for the novel human enzyme polypeptides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9742 BP; 2382 A; 2564 C; 2516 G; 2280 T; 0 U; 0 Other;
 Query Match 9.4%; Score 366.4; DB 4; Length 9742;
 Best Local Similarity 56.3%; Pred. No. 2.7e-65;
 Matches 1043; Conservative 0; Mismatches 631; Indels 178; Gaps 12;
 370 TATTATTAGACAGAGTTCTCACTC-TGTCACCCAGGCTGGAGTGCAGTGGCAATCAT 428
 6351 TTTTCCTTGGATGAGGTTTCACTCTTTGTTGCCAGGCTGGAGTGTGGCGGTATCTT 6392
 429 AGCTCACTGTAACCTCAAACTCTCTGGGCTCAAGTGTACTACTCTCAGCTCCAGAGT 488
 6291 GGCTCATTGCAACCTCCGGCTCCGGGTTCAAGCAATCTCTCTGCTCAGCTCTGAGT 6232
 489 AGCTAGAGCTACAGGCACACAGCCATACCTGCTAAATTTTTTTTTTAAATTTTCAATT 548
 6231 AGCTGGGGTTACAGGCATGTGCCACCGCCGGCTAAATTTTGAATTTTA----- 6181
 549 TATGTAATTCATTTCTTTCTTTTGTGTTGTTTGGATGAGGTCTCACTTTGTT 608
 6180 -----GTAGAGGAGGGTTTCTCATGTT 6157
 609 ACCCAGGCTGAGGGCAGTGGCATGGTGACAGCTGACAGGCTTTGACTTCTCTGGGCTCAA 668

Db 6156 GGTCTAGGCTGGT-----CTTGAACCTCCGACCTCAG 6126
 QY 669 GTGATCCTCCTCCTCAGCCTCCCAAGTAGCTGGGAGTCAACAAACAGCTGTACCATGCT 728
 Db 6125 GTGATCCGCGCGCTCGGCTCCCAAAAGTGTGGGATTACAGCGGTGAGCCACAGCGCTC 6066
 QY 729 GGTGATATTTTTTTTCTTTGAAACAGGTATCACTCTGTGTCGCCAGCTGGAGTACAGTG 788
 Db 6065 GGCCTTTTTTTTTTATTTTGGATGGAGTTTGCCTGTGTCAACCAGACTGGAGTGCAGTG 6006
 QY 789 GGTAAATATAGTCACTGACGCTC-CCCTCCTGGGCTCAAGCAATCCGCTGGCCTCAG 847
 Db 6005 GTGAGTCTCAACTTACTGCAACCTCTGCTGCTAGGTTCAAGCAATCTCATGCTCAG 5946
 QY 848 CATCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGCCGCTCAAGCTTAAGCTTTAA 907
 Db 5945 CTTCCCAAGTAGTGGGACTACAGGCATGTGCCACACCGCTGGCTAAATTTTTTATA 5890
 QY 908 GATTTTTTGTATAGGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTTTATTGTCGAG 967
 Db 5889 TTTTTTTAGTAGAGAGCTGCTGCTGCTGTTGGCTAGGCTGGTCTGGAACGCTCAGCT 5830
 QY 968 AAGGTCTCATATGTTGGCCATGATCCCCCATCTCCACTTCCAAAGTGTCTATCTTATC 1027
 Db 5829 CAAG-----TGATCCGCCCACTCCGCTCCCAAGTGTGGGATTATA 5786
 QY 1028 TGTTCATTAGTCACTGAGACAGATTTAGGTTGTTTCCACTTTTGGACCATATGATTA 1087
 Db 5785 GCGGTGAGCCACGGCATCTGAGTGCATGAACATTTTAAAGGGCCAGTCATGATTAAT 5726
 QY 1088 TACTCCAGTGAATATTCATGTTATACATTTGTGGGCATATGTTTTCATTTCTGTGGGT 1147
 Db 5725 GTCTCTGGGTTAAATTCATGACCGGCTTGGGTTATGACAGAGTTCTAGCATAGCC 5666
 QY 1148 TTATATCTAGGAGTGGAAATGCTGGATCCCGGTATATATTTTGACAGGAGATTGAGG 1207
 Db 5665 CATTTTACAGATAGGACCACTGAGTGGGATCTTAGGGAAT----- 5629
 QY 1208 GAAGAAACTTGGGAAATGAAGCATGTTTGAATACAGCAG-AGTCAGCGGTTTTT 1266
 Db 5628 -TAGAGAAATTAGGAACTTGTAAACGGTGTGACAGCCAGGAGTATAGGAGCTAAGAA 5570
 QY 1267 CGGAGTTTTATTTATTTATTTCTGTGTGACAAATGTGACAGTTTGATGAAGATACAA 1326
 Db 5569 AGCTTTGACGCTAGGCATGGGAGCTCACACCTGTATCTCAGTGTCTTGGAGGCTGAG 5510
 QY 1327 TAAGTGAAGTGAAGTAATTAAGCTGGAATAGGGGCTTCAAGTAAATCATGAGCACT 1386
 Db 5509 GCAGGAGGATCGCTTGGAGCCAGGAGTTTGAGACCAAGCTTGGGGACATAGCGAGACCT 5450
 QY 1387 TTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAATAATCAAAATTTT 1446
 Db 5449 ATCTATGAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5390
 QY 1447 TTTTTTTTTGAAAGAGTCTTGTCTCTTTCACTGCTGGCTGGAGGGGAGTGTGTATCT 1506
 Db 5389 TTTTTTTTTGACAGAAATCTCACTCTGTGTCGCCAGGCTGGAGTGCATGTTGTTATCT 5330
 QY 1507 CAGCTCAGTCAACTTTTGGCTCCGGTTCAGCAATTTCTCTGCTTCAAGCTCCCAAG 1566
 Db 5329 CAGCTCAGCAACCTCCGCTCTCAGGTTCAAGTGAATCTCTCTCCTCAGCTCTCTGAG 5270
 QY 1567 TAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTTTGTATTTTATAGTAGAGA 1626
 Db 5269 TAGTTGGGACTACAGGGGGTGCACCAATGTCGGCTAATTTTTTGTATTTTATAGTAGAGA 5210
 QY 1627 TGGGATTTCACTTTGTGTGGCCAAAGCTGGTCTCAAACTTTTTTGTCTGTCAATTTG 1680
 Db 5209 CGGGGTTTCACTATGTTGACAGGCTGGTCTCGAACTCCTGACCTTGTGATGTGCGGCC 5150
 QY 1681 -----TTGTAATAATTTGTTCC 1696
 Db 5149 TCGGCTCCCAAAAGTGTGGGATTACAGGTGTGAGCCACTCGCTGGCGCTTGTGTTT 5090

QY	1697	TTTTGCTGAGTAGGCCCCCAGACCAAAAATAAACTTTAGATCCAAATCAGTGTG	1756
DB	5089	TTCTTTTACAAACACGCGGCCAGCAGATATCTTATTTCTTTAGAAAACCTAAGTGG	5030
QY	1757	TTGGTTTGACCACTGTCTCACTTGAGAACCAACAGTGTGACCGGCTCAGAGT---	1812
DB	5029	AGATAGACAGACAGTACTCAGATCATTCATATAACCGGTTAGTTGTAATTCACA	4970
QY	1813	GGTATCTCTGTCGAAAGAGAAAATAGAAATATTTCTCCGGCCAGGGGTG--GTG	1870
DB	4969	TTGGATCTGACCTCCAGGATGCACCTGTTAGAAAGTGGCTTGGGCCAGGGGTGTG	4910
QY	1871	GCTCATGCTGTAAATCCAGCACTTTGGGAGCCAGGCGATGTGGATCACCCTGAGT	1930
DB	4909	GCTCACACCTGTAATCTAGCGCTTTGGGAGGCGAGGCGAGCTGATTCCTTGAGT	4850
QY	1931	GAGTTCAAAACCAAGCTGCGCAACATGTGTGAACCCCGTCTCTACTTAAATAACA	1990
DB	4849	TAGTTAATAAATAGCTGTGCACATGTTGAACCTCATCTCTACTTAAATAACA	4791
QY	1991	TTAGCTAAGTGTGGCGCATGCTGTAAATCCAGCTACTTGGGAGGCTGAGCGAGG	2050
DB	4790	TTAGCCAGCGGTGGTGCACCCACCTATATAGTCCAGCTACTTGGGGGGCTGAG	4731
QY	2051	AATTTCTTGAACCGGAGGCGAGGTTGCGTGAAGCGAGATCACACCACTGCATCCA	2110
DB	4730	AACTGCTTGAAGCCGGAGGTGGAGTTGCAGGAGGCCAGATGTGCCACTGCACTTA	4671
QY	2111	GCTTGGGGAGAGGAGGAGACTTCTCTCAAAAAACAAAAACAAAGAAAT	2162
DB	4670	GCTTGGGAGACAGACAGACTCCATCTCAAAAAAAGAAAAGT	4619
RESULT 8			
ID	AAX90847	standard; DNA; 84607 BP.	
XX	AAX90847;		
XX	13-JAN-2000	(first entry)	
DE	Human PACAP	genomic sequence.	
XX	Pituitary adenylate cyclase activating polypeptide; PACAP gene; probe;		
KW	neuroendocrine disorder; bipolar affective disorder; BP-I; UMR;		
KW	severe bipolar affective disorder; schizoaffective disorder manic type;		
KW	SAD-M; chromosome 18p11; retinal ganglion cell; STS marker; ADCYAP1;		
KW	dinucleotide repeat polymorphism marker; W3440; detection; mutation;		
KW	clinical observation; subjective report; compound; mapping; diagnosis;		
KW	treatment; genomic sequence; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	5'UTR	72668..72991	
FT	FT	/*tag= a	
FT	CDS	72993..77243	
FT	FT	/*tag= b	
FT	FT	/*product= "PACAP protein"	
FT	FT	/*note= "includes exons and introns"	
FT	FT	77244..78651	
FT	FT	/*tag= c	
FT	FT	/*note= "Comprises a dinucleotide repeat polymorphism	
FT	FT	marker, W3440 and an STS marker, ADCYAP1"	
XX	WO9951762-A1.		
XX	14-OCT-1999.		
XX	02-APR-1999;	99WO-US007401.	
XX	03-APR-1998;	98US-0080570P.	

XX	(MILL-) MILLENNIUM PHARM INC.		
PA	(REGC) UNIV CALIFORNIA.		
XX	Chen H, Freimer NB;		
PI	WPI; 1999-611046/52.		
XX	Use of a gene linked to the region of a chromosome associated with		
PT	neuroendocrine disorders, especially for diagnosis of bi-polar		
FT	affective disorder and schizoaffective disorder manic type.		
XX	Disclosure; Fig 2A-2U; 99pp; English.		
XX	The present genomic sequence is that of human pituitary adenylate cyclase		
CC	activating polypeptide (PACAP). This gene is localised to the short arm		
CC	of chromosome 18p11 and synthesised in the retinal ganglion cells. The		
CC	genomic DNA is isolated from lymphocytes of individuals. The coding		
CC	region of the gene has a dinucleotide repeat polymorphism marker known as		
CC	W3440 and an STS marker, ADCYAP1 is located in the 3'UTR. These markers		
CC	are useful to map the gene physically and genetically to the chromosome		
CC	18. This sequence can be used to design probes for detecting the presence		
CC	or absence of a mutation in the wild type PACAP gene. It is also useful		
CC	for genetic diagnosis of neuroendocrine disorders like bi-polar		
CC	affective disorder (BP-I) and schizoaffective disorder manic type (SAD-M),		
CC	which are more valid than current clinical diagnosis based on clinical		
CC	observation and subjective reports. It is used for identifying compounds		
CC	that modulate PACAP gene expression which are useful for treatment of		
CC	PACAP disorders		
XX	Sequence 84607 BP; 23055 A; 17765 C; 18980 G; 24744 T; 0 U; 63 Other;		
SQ	Query Match 9.1%; Score 356.2; DB 2; Length 84607;		
	Best Local Similarity 68.6%; Pred. No. 5.5e-63;		
	Matches 522; Conservative 0; Mismatches 233; Indels 6; Gaps 2;		
QY	1429 AAAAAATCAAAATTTTTTTTTTTTTCAGAAAGAGTCTTCTTTTCCCTGCTGG 1488		
DB	20031 ACAACTTTATTTATTTATTTATTTATTTGAGAGGAGTCTGCTCTGTTCCTGCTG 20090		
QY	1489 AGGCGAGTGTGTGATCTCAGCTCACTGCACTTCGCTCCGCGGTTCAAGCAATTC 1548		
DB	20091 ACTACAAAGCGATGTTCTCAGCTCACTGCACTTCGCTCCGCGGTTCAAGCAATTC 20150		
QY	1549 CTGCTTACGCTCCCAAGTAGCTGGGACTACAGGCACTTCCCACTGCCCAGCTGATTT 1608		
DB	20151 ATGCTTACGCTCCGAGTAGCTGGGATTCAGGCGTGGCCACCAACACCACTAATTT 20210		
QY	1609 TTGTATTTTATGAGATGGGATTTTCTTTTGGGCAAGTCTGCTTCAAACTTTTGG 1668		
DB	20211 TTGTATTTTATGAGATTTGGGTTTTCACCATGTTGGCCAGGCTGCTCTCGAAC---TC 20266		
QY	1669 CTGTCATATTTGTTAACTATTGTTCTTTGCTGAGTGGGCCCCCAGACCAAAAA 1728		
DB	20267 CTGACCTCGTGAAGTTCGCCATCTCGGCTCTCTAAAGTGTGGATTCAGGTTGAGTCA 20326		
QY	1729 AATAAATCTTAGAATCCAAATCAGTGTGTGTTTGACCACTGTCTCACTTGAACACACAG 1788		
DB	20327 CCGTGCCTGCCTAACAAATTGATTAGACGGTGGATAGATGATTATTTATTGCTGTGT 20386		
QY	1789 TGTGACCGGCTCAGGAGTGAAGTGTCTGCTCGAAAGAGAAATAGATGAAT 1848		
DB	20387 TCTAGTTGTATCCAAAAAATCAAAATAAGAAATTAACCAAAATAGGACACATTTAAGAA 20446		
QY	1849 ATCTCCGGCCAGCGGTGTGCTCATCTCCCTGATTCAGCACTTTGGGAGGCCAAGG 1908		
DB	20447 ATCATACGGCGCGGCGGTGCTCACGCTGTAATCCAGCACTTTGGAGGCCGAGG 20506		
QY	1909 CATGTGATCAGCTGAGGTCAGGAGTTCAAAACAGCTGGCCAAACATGTTGAACCCCG 1968		
DB	20507 TGGGCGGATCAC--GAGGTTCAGGAGATCGAGACCATCTCTGGCTAAACCGTGAACCCCG 20564		
QY	1969 TCTCTACTAAAAATCAAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTATATCCAGCT 2028		

```
Db      20565 TCTCTACTAAAAATACAAAAATAGCCGGCGCTAGTGGCGGCGCTGTCGTCCTCCAGCT 20624
QY      2029  ACTTGGAGGGTGAGCAGAGAAATTTCTTGAACCGGAGGCGAGAGTTGCAGTGAAGC 2088
Db      20625 ACTCGGAGGCTGAGCAGAGAAATCGCGTGAACCGGAGGCGGAGCTTGCAGTGAGCC 20684
QY      2089  GAGATCACACCACTGCACCTCAGCCTCGGGGAGAGAGCGAGACTTCTCTCAAAAAACA 2148
Db      20685 AAGATCGGCGCACTGCCCTCCAGCCTGGCGCAGAGCGAGACTCCGCTCTCAAAAAAAA 20744
QY      2149  AAAAAAAGAAATTAAGCAATTAACATTCAGAGAGAA 2189
Db      20745 AAAAAAAGAAATTAAGCAATTAACATTCAGAGAGAA 20785

RESULT 9
ID      AAD14749 standard; DNA; 12394 BP.
XX      AC      AAD14749;
XX      DE      01-NOV-2001 (first entry)
XX      KW      Human glycogen synthase kinase 3 alpha genomic DNA.
XX      DE      Human; glycogen synthase kinase 3 alpha; antidiabetic; cytostatic;
KW      antisense therapy; diabetes; hyperproliferative disorder; inflammation;
KW      neurological disorder; tumour; haematopoietic disorder; infection;
XX      hyperproliferative disorder; developmental disorder; ds.
OS      Homo sapiens.
XX      FH      Key
XX      FT      CDS
XX      FT      Location/Qualifiers
XX      FT      115..11786
XX      FT      /tag= a
XX      FT      /product= "Human glycogen synthase kinase 3 alpha"
XX      FT      115..397
XX      FT      /tag= b
XX      FT      intron 398..2437
XX      FT      /tag= c
XX      FT      /tag= C
XX      FT      2438..2625
XX      FT      /tag= d
XX      FT      2626..5638
XX      FT      /tag= e
XX      FT      exon 5639..5722
XX      FT      /tag= f
XX      FT      intron 5723..5863
XX      FT      /tag= g
XX      FT      exon 5864..5974
XX      FT      /tag= h
XX      FT      intron 5975..7901
XX      FT      /tag= i
XX      FT      exon 7902..8032
XX      FT      /tag= j
XX      FT      intron 8033..8120
XX      FT      /tag= k
XX      FT      exon 8121..8227
XX      FT      /tag= l
XX      FT      intron 8228..9196
XX      FT      /tag= m
XX      FT      exon 9197..9294
XX      FT      /tag= n
XX      FT      intron 9295..9374
XX      FT      /tag= o
XX      FT      exon 9375..9470
XX      FT      /tag= p
XX      FT      intron 9471..9897
XX      FT      /tag= q
XX      FT      exon 9898..10084
XX      FT      /tag= r
XX      FT      intron 10085..10430
XX      FT      /tag= s
```

```
FT      exon 10431..10523
FT      /tag= t
FT      intron 10524..11712
FT      /tag= u
FT      exon 11713..11786
FT      /tag= v
XX      WO200152865-A1.
XX      PN      26-JUL-2001.
XX      PF      16-JAN-2001; 2001WO-US001411.
XX      PR      21-JAN-2000; 2000US-00488856.
XX      PA      (ISIS-) ISIS PHARM INC.
XX      PI      Monia BP, McKay R, Butler MM, Wyatt JR;
XX      WI      WPI; 2001-442247/47.
XX      DR      P-PSDB; AAE06521.
XX      PT      Antisense compound 8 to 30 nucleobases in length comprising a compound
XX      PT      that is targeted to a nucleic acid molecule encoding glycogen synthase
XX      PT      kinase 3 alpha, useful for the treatment of e.g. diabetes and
XX      PT      hyperproliferative disorders.
XX      PS      Example 15; Page 92-100; 115pp; English.
XX      CC      The invention relates to an antisense compound 8 to 30 nucleobases in
XX      CC      length targeted to a nucleic acid encoding glycogen synthase kinase 3
XX      CC      alpha. The antisense compound specifically hybridises with and inhibits
XX      CC      the expression of glycogen synthase kinase 3 alpha. The antisense
XX      CC      compound is useful for the treatment of a diseases associated with
XX      CC      glycogen synthase kinase 3 alpha such as diabetes, a neurological
XX      CC      disorder, a haematopoietic disorder, a hyperproliferative disorder or a
XX      CC      developmental disorder. The antisense compounds may also be used
XX      CC      prophylactically to prevent or delay infection, inflammation or tumour
XX      CC      formation. The present genomic DNA sequence encodes human glycogen
XX      CC      synthase kinase 3 alpha protein
XX      SQ      Sequence 12394 BP; 2780 A; 3043 C; 3547 G; 3024 T; 0 U; 0 Other;
XX      Query Match 8.9%; Score 349.8; DB 4; Length 12394;
XX      Best Local Similarity 55.2%; Pred. No. 7.3e-62;
XX      Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;
QY      370 TATTATTAGACAGGTTCTCACTCTGTCCACCCAGGCTGGAGTGCAGTGGCAATCATA 429
Db      3604 TTTTITTTTGGATGAGTCTTACTCTGTCCACCCAGGCTGGAGTGCAGTGGCAATCTTG 3663
QY      430 GTCCTACTGTAACCTCAAACTCCTGGGCTCAAGTATFCTTACTACTCCTCAGCCTCCAGAGTA 489
Db      3664 GCCCGCTACAACTCTGCTCCAGGTTCAAGCGATTCTTGTGCCTCAACCCACCTGAGTA 3723
QY      490 GCTAGGACTACAGGCACACACAGCCATACCTGCGTAATTTTTTTTTTTTAAATTCATTTT 549
Db      3724 GCTGGATTGCGAGCATGGGCCACATGCTGGCTGAATTTTTTTTTTTTGTATT--TTTA 3781
QY      550 ATGTATTCATTTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Db      3782 GTAGAGACAGGGTTTGGCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTAAT 3841
QY      610 CCCAGGCTGGAGGGGAGTGGGATGGTGACAGCTGAGCAGCCCTTGACTTCTCTGGGCTCAAG 669
Db      3842 CCGCCCCACCTCGACCTCCCAAAGTGTGGTATTATACAGGCATGA-----GCCAGCAT 3894
QY      670 TGATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGCTGTCA--CCATGCGCT 728
Db      3895 ACCTAGCCCTGATTTTTCAGGAACTGAAACTGAAATTTAGATGTGAAATCTTTTTT 3954
QY      729 GGCCTGATATTTTCTTTGAAACAGGGTATCACTCTGTGTGCCAGGCTGGAGTACAGTG 788
```


Db 3955 TTTT...TTTGTGAGCGAGTCTCATGCTGTCACCCAGGCTGGAGTGTGGTG 4014
Qy 789 GCGTAATAATAGTCACTGAGAGCTCC-CTTCTGGGCTCAAGCAATTCGCTGCGCTCAG 847
Db 4015 GCGTGAATCTCGGTCCTGCAACCTCCGCTGCGGGTTCAGCGGATCTTCTGCTCAG 4074
Qy 848 CATCTCAGTAGCTGGGACTACAGCTTGTGCCACACAGGCGCCAGCTTATTTTAAAAAT 907
Db 4075 CCTCCCTAGTAGCTGGGACTACAGCGTGTGCCACCACTCTCGGCTAATTTTGTATTT 4134
Qy 908 GATTTTGGTATAGAGGAGGCTTGTCTATGTGTCTCAGGCTGTATTTTATTTGTTG--- 963
Db 4135 CTAGTAGGAGCGGGTTTCAACCGAGTTAGCCAGGATGCTCTATTTTATTTTATTTT 4194
Qy 964 ----AGACAAGGTTCTACTATGTTGCCATGATCCCCACCTCCACTTCCCAAGTGCTC 1019
Db 4195 TTTAAGACAGAATCTCGTCTGTCTACTAAGGCTGGAGTGCAGTGTGTGATGTGGCTCA 4254
Qy 1020 ATCTTATCTGTTTCATTAGTTCAGTTCAGACAGATTTAGGTGTGTTTCCACTTTTGTACCAAT 1079
Db 4255 CTGCAACCTCTGCTCTCCCTGGGTTCAAGCGCTGCAACCTCTGCCCTCTGGGTTCAAGCAAT 4314
Qy 1080 ATGAATAATACCTCAGTAGTAATATCATGTATACATTTGTGTGGGCAATGTTTTCATTTC 1139
Db 4315 TCTTGTACCTCATCCACCTGAGTAGTTGGAATCACAGCGGTGGCCACCATGCTGCCAGCTA 4374
Qy 1140 TGTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGGTAAATTTTGTACAGGCAGA 1199
Db 4375 ATTTT...TTTGTATTTTGTAGAGATGGGGTTTTCACGTTGGCCAGGCTGGTCTCGAAC 4434
Qy 1200 GTTCAGGGAAGAAAA-----ACPTGGGAAAAATGAAGCATGTTTGAANAATCA 1246
Db 4435 TCCTGGCTCAAGTGATTTCCCTGCTTGGCTCCCAAGTCTGGGATTCAGGGGTGA 4494
Qy 1247 GCAAGAGTGCAGGGTTTTCGGAGTTTATTTATTTCTGTTGACAAAATGTCAGTTT 1306
Db 4495 GCCACATGCCAGCTGTTTATTTATTTTATTTTATTTTAAAG--GCTGGGTATGTAGCT 4552
Qy 1307 GATGAAGATCAAGTTTACTTAAGTGAAGAGTGAATTTAAGGCTGGAATAGGGCTTCA 1366
Db 4553 CATGCTGTAACTCTTGAACTTTGGAGGCC-CAAGGAGGAGATTTGCTGAGACTAGA 4611
Qy 1367 GAGTAAATCATGAAGCACTTTGAATACCAAAATTAAGGAGTGTGCTGTAAACAAATA 1426
Db 4612 GTTCAARACCACTGGCCACATAGCCAGGTCTTTTAAAAATAATAATAATAAAT 4671
Qy 1427 ATAAAAATCAAAATTTT...TTTGTGAAAGAGTCTTGTCTTTTCAACCTGGCT 1486
Db 4672 TTTATCTTATTTATTTATTTATTTATTTATTTTGTGACAGAGTCTGTCGCCAGGCT 4731
Qy 1487 GGAGGCGAGTGTGTGATCTCAGCTCAGCTGCACTTTCGCCCTCCGGGTTCAAGCAATTC 1546
Db 4732 GAGGTGAGTGGCGGATCTCAGCTCAGCTGCAAGCTCCGCTCTGGGTTCAAGCAATTC 4791
Qy 1547 TCTGCTTCAGCTCCCAAGTGTGGGACTCAGGCACTTCCACCATGCCAGCTGA- 1605
Db 4792 TCTGCTCAGCTCCGAGTAGCTGGAGTCTCAGGCTGCTGCAAGTGTGGCTTAAAT 4851
Qy 1606 ----TTTGTATTTTGTAGTAGAGTGGATTTCACTTTGTGGCAGCTGTTCTCAACT 1663
Db 4852 TTTT...TTTGTATTTTGTAGAGCAGGGTTTCAACCGTTTAAACAGGATGGTCTCAATCT 4911
Qy 1664 TTTTGTGTCAATAATTTTGTAACTATTGTTCTTTTGTGAGTGGGCCCCAGACCA 1723
Db 4912 CTTGAC-----TTCTGTATCCACCACTCAGCTCCCAAGTGTGGGATTCAGGCGT 4966
Qy 1724 AAAAAAATAATCTTGAATTCAAATCAGTGTGTGGTTTGGACCTGTCACCTTGAGAAC 1783
Db 4967 GAGCCACCGCTGGCCCT-----GGCCTATCCTTTTAAACCTTTATTTTGGAGAA 5019
Qy 1784 CACAGTGTACCGGCGCTCAGAGTAGAGGTGATCTCTGCTGAAAGAGAAATAGAG 1843
Db 5020 AAAAAATCAGAAGGTGCCATTTGGCTTTTACATGTCAGCAATTAAGTTGAAAAAATTTT 5079

Qy 1844 AAAATATTTCTCCGGGCGCAGGCTGGTGGCTCATGCTGTAAATCCACGACTTTTGGAGGC 1903
Db 5080 TTTT...TAAAGTGGGCTGGCTGGGCGGGTGCCTCAGCCCTGTAAATCCACGACTTTTGGAGGC 5139
Qy 1904 CAAGGCAATGTGGATCA-----CCTGAGGTGAGGAGT 1934
Db 5140 TGAGGCTGTGGATCATGAGTCAAGGAGGCTGAGGAGGTGATCAACAGGTGAGGAG 5199
Qy 1935 TCAAAACAGCCTGGGCAACATGCTGAAACCCCGTCTCTACTATAAAATACAAAAATTAG 1994
Db 5200 TCGAGACCACTCTGGCTAACTGCTGAAACCCCATCTCTACTATAAAATAC-AAAAATTAG 5258
Qy 1995 CTAAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTGGAGGCTGAGGAGGAGTAT 2054
Db 5259 CTGGGCGTGGTGGTGCATGCTGTAAATCCCAAGCTACTTGGAAAGGCTGAGGAGGAGTAT 5318
Qy 2055 TCTTGAACCCGGGAGGAGGAGTGTGAGTGAAGCGAGATCACACACTGCACTCCAGCT 2114
Db 5319 GCTTGAACCCGGGAGGAGGAGTGTGAGTGAAGCGAGATTTGAGGAGTGCATGCCAGCT 5378
Qy 2115 GGGGAGAGAGGAGAGACTTCTCTCAAAAAACAAAAA 2153
Db 5379 -GGCAACGAGCAAGACTCTCTCAAAAAA 5416

RESULT 10
AAL05122
ID AAL05122 standard; DNA; 13216 BP.

XX AAL05122;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 7810.

XX Human; reproductive system related antigen; reproductive system disorder; cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	XX	(HUMA-) HUMAN GENOME SCI INC.	XX	Sequence 13216 BP; 3041 A; 3400 C; 3243 G; 3532 T; 0 U; 0 Other;
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.	XX		XX	
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.	XX		XX	
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.	XX		XX	
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.	XX		XX	
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.	XX		XX	
PR	22-AUG-2000;	2000US-0226581P.	PR	17-NOV-2000;	2000US-0249207P.	XX		XX	
PR	22-AUG-2000;	2000US-0226586P.	PR	17-NOV-2000;	2000US-0249208P.	XX		XX	
PR	22-AUG-2000;	2000US-0226586P.	PR	17-NOV-2000;	2000US-0249209P.	XX		XX	
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.	XX		XX	
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249211P.	XX		XX	
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249212P.	XX		XX	
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249213P.	XX		XX	
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249214P.	XX		XX	
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249215P.	XX		XX	
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249216P.	XX		XX	
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249217P.	XX		XX	
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249218P.	XX		XX	
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249244P.	XX		XX	
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249245P.	XX		XX	
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249264P.	XX		XX	
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249265P.	XX		XX	
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249297P.	XX		XX	
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.	XX		XX	
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249300P.	XX		XX	
PR	08-SEP-2000;	2000US-0232080P.	PR	01-DEC-2000;	2000US-0250160P.	XX		XX	
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.	XX		XX	
PR	12-SEP-2000;	2000US-0231968P.	PR	05-DEC-2000;	2000US-0251030P.	XX		XX	
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251988P.	XX		XX	
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0256719P.	XX		XX	
PR	14-SEP-2000;	2000US-0232399P.	PR	06-DEC-2000;	2000US-0251479P.	XX		XX	
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251856P.	XX		XX	
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251868P.	XX		XX	
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251869P.	XX		XX	
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.	XX		XX	
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.	XX		XX	
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.	XX		XX	
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.	XX		XX	
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2000US-0255678P.	XX		XX	
PR	25-SEP-2000;	2000US-023499							

Qy 1621 TAGAGATGGGATTTACATTTGTTGGCCAGCTGGTCTCAAACTTTTGTGTCATAATTG 1680
Db 3805 TACAGATGGGGTTTACCACATGTGTCTAGGCTGGTCTGAACTCTGACTTCATGATCCG 3864
Qy 1681 TTGTAACATATTGTTCTCTTTGCTGAGGTAGGGCCCCAGACCAAAAAATAAATCTTAG 1740
Db 3865 TCCCCCTCGACTCCCAAGTGTAGGATTACAGGCATGAGCACTGTGCCAGCC--- 3920
Qy 1741 AATCAAAATCAGTGTGTTGGTTGACCACTGTCATTGAGAACCAACAGTGTGACCAAGGC 1800
Db 3921 ---GCTAAATTTATTTTGTACAGACAGAGTTTCTCCATGTTGCCAGGCTGATCTGAA 3977
Qy 1801 CTCAGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGCC 1860
Db 3978 CTCCTGGGCGCAAGTGAACCACTGCTCGGC---CTCAAGTGTGGGATTACAGGCC 4034
Qy 1861 AGGCGTGGTGGTCTATGCTCTGCTGTAATCCAGCACTTTGGAGGCCAAGGCATGATCAC 1920
Db 4035 GGGTGGGTGGTCTACGCTGTAATCCAGCACTTTGGAGGCCAAGGCATGATCAC 4094
Qy 1921 CTGAGGTACAGAGTTCAAAACAGCTGGCCCAACATGTTGAAACCCGCTCTACTATAAA 1980
Db 4095 CTGAGGTCCGGAGTTCAAGACAGAGCTGGCCCAACAGTGTGAAACCCGCTCTACTATAAA 4154
Qy 1981 ATACAAAAATTAGCTAAGTGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGT 2040
Db 4155 ATAC-AAAAATTAGCACTGTTGGTGGCGCATGCTGTAATCCAGCTACTTTGAGAGGT 4213
Qy 2041 GAGGCAGGAGAAATTTCTGAAACCCGGAGCGAGAGTTGCAAGCGAGATCACCA 2100
Db 4214 GAGGCAGGAGAAATCACTTGAACCCAGGAGCGAGAGTTGCAAGCGAGATCACCA 4273
Qy 2101 CTGCACTCCAGCTGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAGA 2160
Db 4274 CTGCACTCCAGCTGGGGAGAGCGAGAGTGTCTCTCAAAAAACAAAAACAAAGA 4333
Qy 2161 ATTAAGCAAAATTAGACA 2177
Db 4334 ACAAAACAAACAAACA 4350
RESULT 11
ID ABL98014 standard; DNA; 13216 BP.
XX AC ABL98014;
XX DT 21-JUN-2002 (first entry)
XX DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2656.
XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX OS Homo sapiens.
XX WO200155317-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001329.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

PR	14-SEP-2000;	2000US-02332393P.	PR	05-DEC-2000;	2000US-0256719P.	Query Match	8.9%;	Score 347.4;	DB 4;	Length 29163;	
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.	Best Local Similarity	69.9%;	Fred. No. 2.8e-61;			
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.	Matches	515;	Conservative	0;	Mismatches	211;
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251869P.					Indels	11;
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.					Gaps	3;
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.						
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.						
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.						
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.						
PR	25-SEP-2000;	2000US-0234998P.	XX	(HUMA-) HUMAN GENOME SCI INC.							
PR	26-SEP-2000;	2000US-0235484P.	XX	Rosen CA, Barash SC, Ruben SM;							
PR	27-SEP-2000;	2000US-0235834P.	XX	WPI; 2001-465570/50.							
PR	27-SEP-2000;	2000US-0235836P.	DR								
PR	23-SEP-2000;	2000US-0236367P.	XX								
PR	29-SEP-2000;	2000US-0236368P.	XX								
PR	29-SEP-2000;	2000US-0236369P.	XX								
PR	29-SEP-2000;	2000US-0236370P.	PT	Isolated nucleic acid molecule encoding a reproductive system antigen is							
PR	02-OCT-2000;	2000US-0236802P.	PT	used in preventing, treating or ameliorating a medical condition.							
PR	02-OCT-2000;	2000US-0237037P.	XX								
PR	02-OCT-2000;	2000US-0237038P.	PS	Disclosure; SEQ ID NO 7809; 1297pp + Sequence Listing; English.							
PR	02-OCT-2000;	2000US-0237039P.	XX								
PR	02-OCT-2000;	2000US-0237040P.	CC	The present invention provides the protein and coding sequences of a							
PR	13-OCT-2000;	2000US-0239935P.	CC	number of human reproductive system related antigens. These can be used							
PR	13-OCT-2000;	2000US-0239937P.	CC	in the prevention and treatment of reproductive system disorders,							
PR	20-OCT-2000;	2000US-0240960P.	CC	including cancer. The present sequence is a genomic sequence encoding a							
PR	20-OCT-2000;	2000US-0241826P.	CC	protein of the invention							
PR	01-NOV-2000;	2000US-0244517P.	XX								
PR	08-NOV-2000;	2000US-0246474P.	SQ	Sequence 29163 BP; 6801 A; 7298 C; 7317 G; 7747 T; 0 U; 0 Other;							
PR	08-NOV-2000;	2000US-0246475P.									
PR	08-NOV-2000;	2000US-0246476P.									
PR	08-NOV-2000;	2000US-0246477P.									
PR	08-NOV-2000;	2000US-0246478P.									
PR	08-NOV-2000;	2000US-0246523P.									
PR	08-NOV-2000;	2000US-0246524P.									
PR	08-NOV-2000;	2000US-0246525P.									
PR	08-NOV-2000;	2000US-0246526P.									
PR	08-NOV-2000;	2000US-0246527P.									
PR	08-NOV-2000;	2000US-0246528P.									
PR	08-NOV-2000;	2000US-0246532P.									
PR	08-NOV-2000;	2000US-0246609P.									
PR	08-NOV-2000;	2000US-0246610P.									
PR	08-NOV-2000;	2000US-0246611P.									
PR	08-NOV-2000;	2000US-0246613P.									
PR	17-NOV-2000;	2000US-0249207P.</									

Qy	2041	GAGGAGGAGAAATTTCTTGAACCCGGGAGGAGAGGTTTGCAGTGAAGCGAGATCACACCA	2100
Db	4212	GAGGAGGAGAAATTTCTTGAACCCGGGAGGAGAGGTTTGCAGTGAAGCGAGATCACACCA	4271
Qy	2101	CTGCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA	2160
Db	4272	CTGCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA	4331
Qy	2161	ATTAAGCAAAATTAGACA	2177
Db	4332	ACAAAACAAACAAAAACA	4348

RESULT 14
 ID ABL98013 standard; DNA; 29163 BP.
 XX
 AC ABL98013;
 XX
 DT 21-JUN-2002 (first entry)
 XX
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2665.
 XX
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disease; infection; cytostatic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200155317-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001329.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.

PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	08-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	12-SEP-2000;	2000US-0232081P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	25-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.

PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 2665; 765pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention
XX
SQ Sequence 29163 BP; 6801 A; 7298 C; 7317 G; 7747 T; 0 U; 0 Other;
Query Match 8.9%; Score 347.4; DB 4; Length 29163;
Best Local Similarity 69.9%; Pred. No. 2.8e-61;
Matches 515; Conservative 0; Mismatches 211; Indels 11; Gaps 3;
QY 1441 TTTTCTTTTCTTTTGTGAGAGAGTCTTCTTTCACCTGGCTGGAGGGGAGTGGTG 1500
DB 3623 TCTTTTCTTTTCTTTTCTGAGTCTTGTGCGCCAGGCTGAGTGGAGTGGTG 3682
QY 1501 TGAATCTAGCTACTGCACTTCGGCTCCCGGTTCAAGCAATTCCTGCTTACGCT 1560
DB 3683 CGATCTCAGCTACTGCAACTCCACTCCCGGTTCAAGAGATTCCTCTGCTTAGCT 3742
QY 1561 CCCAAGTAGCTGGGACTACAGGCACCTTCCACCACATGCCAGCTGATTTTGTATTTTAG 1620
DB 3743 CCTGAGTAGCTGGATTACAGGGGTGCACTACCGTGCCTGGCTAAATTTTATATTTTAA 3802
QY 1621 TAGAGATGGATTTCATTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTGTCATATTTG 1680
DB 3803 TACAGATGGGGTTTCCACCATGTTGGTCAGGCTGGTCTCGAACTCTCCTGACTTCATGTCG 3862
QY 1681 TTGTAATATTGTTCTTTTGTCTGAGTAGGGCCCCCAGACCAAAAATAAATCTTAG 1740
DB 3863 TCCCCCTCGACCTCCCAAGATGTAGATTACAGGCATGAGCCACGTGCCCCAGCC --- 3918
QY 1741 AATCCAAATCAGTGTGTGTGTTTGTAGCACTGTCTTGAACCAACACAGTGTGACCCAGGCG 1800

DB 3919 ---GCTAATTTATTTTGTGTACAGACAGATTCTCCATGTTGCCAGGCTGATCTGGAA 3975
QY 1801 CTCAGGATGAGGGTGAATCTCTGCTCGAAAAGAGAAATGAAATATTTCTCCGGGCC 1860
DB 3976 CTCCTGGGCGCAAGTGAACCATCTGCTCGGC---CTCCAAAGTCTGGGATTACAGGCC 4032
QY 1861 AGGCGTGTGCTCATGCTGTAAATCCAGACACTTTTGGAGGCCCAAGCATTGTGGATCAC 1920
DB 4033 GGGTGGTGGTCTCAGGCTGTAAATTCAGCACTTTTGGAGGCCCAAGCAGGTGATCAG 4092
QY 1921 CTGAGGTGAGGATTTCAAAACCCAGCCTGGCCAAACATGTTGAAACCCCTCTCTACTAAA 1980
DB 4093 CTGAGGTCCGAGTTCAAGACCCAGCCTGGCCAAAGGTGACACCTGTCTCTACTAAA 4152
QY 1981 ATACAAAAATTAAGTAAAGTGTGGGCGCATGCTGTAAATCCAGCTACTTTGGAGGGT 2040
DB 4153 ATAC-AAAAATAAGCCACGTGTGTGGGCGCATGCTGTAAATCCAGCTACTTTGAGGGT 4211
QY 2041 GAGGCGAGGAAATTTTGTAAACCCGAGGCGAGGTTGCAGTGAAGCGAGATCACCA 2100
DB 4212 GAGGCGAGGAAATCACTTTGAACCCAGGAGGCGAGGTTGCAGTGAAGCGAGATCACCA 4271
QY 2101 CTGCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGA 2160
DB 4272 CTGCACTCCAGCTGGGCGAGAGAGTGGGCTGTCTCTCAAAAAACAAAAACAAAA 4331
QY 2161 ATTAAGCAAAATTAGACA 2177
DB 4332 ACAAACCAACAAAAACA 4348
RESULT 15
AAS42014/c
ID AAS42014 standard; DNA; 6437 BP.
XX
AC AAS42014;
XX
DT 17-DEC-2001 (first entry)
DE Genomic sequence #330 encoding novel human enzyme polypeptide.
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ds.
OS Homo sapiens.
XX WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001239.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227003P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229503P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 29-SEP-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249277P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251930P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-02559678P.
 XX (HUMA-) HUMAN GENOME SCT INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465566/50.
 XX
 XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 XX treating neural, immune system, muscular, reproductive, pulmonary,
 XX cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX
 XX Disclosure; SEQ ID NO 2140; 1180pp; English.

XX The present invention relates to the isolation of novel human enzyme
 XX polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 XX encoding them. The enzyme polypeptides of the invention may comprise the
 XX functional classes of oxidoreductases, transferases, hydrolases, lyases,
 XX isomerases or ligases. The sequences of the invention are useful in the
 XX diagnosis, treatment, prevention and/or prognosis of a wide range of
 XX disorders including hyperproliferative disorders (e.g. cancer),
 XX immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 XX arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 XX disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 XX cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 XX (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 XX infectious disorders (e.g. Influenza). The polynucleotides of the

CC invention can also be used in gene therapy. AAS41685-AAS42192 represent
CC DNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 6437 BP; 1690 A; 1641 C; 1586 G; 1520 T; 0 U; 0 Other;

```
Query Match      8.9%; Score 347; DB 4; Length 6437;
Best Local Similarity 56.7%; Pred. No. 2.4e-61;
Matches 1052; Conservative 0; Mismatches 625; Indels 180; Gaps 15;

QY 365 TATCATATATTAGAGACAGGTTCTCACATC-NGTCACCCAGGCTGGAGTGGACAC 423
DB 3050 TTTTCTTTCTTTGAGATGGAGTTTCACTCTGTTGTCGCCAGGCTGGAGTGGCGGT 2991

QY 424 ATCATAGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCCTACTACCTCAGCCTCC 483
DB 2990 ATCTTGGCTCAATGCAACCTCCGCTCCCGGTTCAAGCAATCTCCTTGCTCAGCCTCC 2931

QY 484 AGAGTAGCTAGGACTACAGGACACACAGCCATACTGGCTAAATTTTTTTTTTTTAAATTT 543
DB 2930 TGAGTAGCTGGGTTACAGGCATGTGCCACCGCCGCTAATTTGTATTTTAA----- 2875

QY 544 CATTTTAATGATTCATTTTCTTTCTTTTGTGTTGTTGTTGATAGAGGTCTCACT 603
DB 2874 -----GTAGAGGAGGGTTTCTCC 2856

QY 604 TTGTTACCCAGGCTGGAGGCGAGTGGCATGTTGACAGCTGAGCAGCTTGACTTCTGGG 663
DB 2855 ATGTTGGTCAGGCTGGT-----CTTGAACCTCCGAC 2825

QY 664 CTCAGTAGATCTCTGCTCAGCTCAGCTCCCAAGTAGCTGGGACTCAACAACAGTGTACCA 723
DB 2824 CTCAGGTGATCCGCGCTCCGCTTCCCAAGTGTCTGGGATTCACAGGCGTGGACACAG 2765

QY 724 TGCTCGGCTGATATTTTTTTCTTGAACACAGGATACACTGTTGTCGCCAGGCTGGAGTA 783
DB 2764 CGCTCGGCTTTTATTTTATTTTGTAGATGGAGTTTCGCTCTGTACCCAGACTGGAGTG 2705

QY 784 CAGTGGCGTAATAATAGCTCACTGCAGGCCTC-CCCTCCTGGGCTCAAGCAATCCGCTGGC 842
DB 2704 CAGTGGTGAGTCTCAACTTACTGCAACCTCTCGGTCCTAGGTTCAAGCAATCTCATGC 2645

QY 843 CTCAGATCCTGAGTAGTGGACTACAGCTGTGTCACACAGGCCCCAGCTAAGTTTAA 902
DB 2644 CTCAGCTTCCCAAGTAGATGGGACTACAGGCATGTGCCACCAAGCTGGCTAATTTTAT 2585

QY 903 AAAATGATTTTGGTATAGAGAGGTCCTTGCTATGTTGCTCAGGCTGTATTTTATTTGTT 962
DB 2584 A-----TTTTTTTAGTAGAGCGTGGTTTCGCTCTGTTGGCTAGGCTGGTCTGNAAGCGCT 2529

QY 963 GAGACAGGCTCTCATATGTTGCCATGATCCGCCACCTCCACTCCAGTCCCAAGTGTCAATC 1022
DB 2528 GACCTCAAG-----TGATCCGCCACCTCCGCTCCCAAGTGTGGGA 2485

QY 1023 TTATCTGTTCAATTAGTCAATGACACATTTAGGTTGTTTCCACTTTTGTGACCAATATG 1082
DB 2484 TTATAGGCTGAGCCACCGGACTGTAGTGCAATGAACTATTTAAGGCCAGCTCATTTGATC 2425

QY 1083 AATAATACCTCCAGTGAATATTCAATGATATATTTGTTGGGCATATGTTTCAATTTCTGT 1142
DB 2424 TAAATCTCTGTGGGTTAATTTCACTACGCCCTTGGGTTTATGACAGAGGTTCTAGCCA 2365

QY 1143 TGGGTTTATATAGAGGTGGAAATGCTGGATCCCGGTAATATTTTGTAGGAGAGATT 1202
DB 2364 TAGCCCAATTTTACAGATAGGACCAAGTGAAGATCTTAGGGAAT----- 2323

QY 1203 CAGGGAAGAAACCTTGGGAAATGAAGCATGTTTAGAAATCAGCAAG-AGTGCAGGGG 1261
DB 2322 -----TAGAGAAATTTAGGGAACCTTCTAAGCGTTGCAAGCCAGGAAGTATAGGCAGCT 2269

QY 1262 TTTTTCGGAGTTTATTTTATATTTCTGTTTGAACAAATGTGCAGTTTGTATGAAGATCAAGT 1321
```

Search completed: May 12, 2004, 19:47:27
Job time : 2111.64 secs

```
DB 2258 AAGAAAAGTCTTGCAGCTTAGGCACTGGGAGCTCACACCTGTATCTCAGTGTCTTGGAGG 2209
QY 1322 TATACCTAGTGAGAGTGAGAAATTAAGCTGGAAATAGGGCGTTAGAGTAAATCATGAA 1381
DB 2208 CTGAGGCAGGAGGATCGCTTGAGGCCAGAGTGTGAGACCAGCTTGGGAAACATAGCGAG 2149
QY 1382 GCACCTTTGAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATAAAAAATCACAAT 1441
DB 2148 ACCCTATCTCTATGAAATAATAACAAATAAATAAGTGTAGAAATATCTTGCTTTTCTT 2089
QY 1442 TTTTCTTTTCTTTTGTGAAAAGAGCTTGCTCTTTTACCTCGCTGGAGGAGGAGTGGTGT 1501
DB 2088 TTTCTTTTCTTTTGTGAGACAGATCTCCTCTGTTGCCAGGCTGGAGTGAATGGTGT 2029
QY 1502 GATCTCAGCTCACTGCAACTTTTGGCTCCCGGTTCAAGCAATTTCTCTGTTTCAAGCTC 1561
DB 2028 AATCTCAGCTCACCAACCTCCGCTCTCAGGTTCAAGTGAATCTCTCCTCCCTCAGCCTC 1969
QY 1562 CCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCCCAGCTGATTTTGTATTTTATGT 1621
DB 1968 CTGAGTAGTTGGGACTACAGGCGGTGCCCATGTCCGCTTAATTTTGTATTTTATGT 1909
QY 1622 AGAGATGGGATTTCACTTTTGTGGCCAAAGCTGGTCTCAAACTTTT----- 1666
DB 1908 AGAGACGGGTTTCACTATGTTGACAGGCTGGTCTCGAATCTCCTGACCTGTGATCTGC 1849
QY 1667 -----TGTGTCAATAATGTTGTGTA-----CTATGTG 1692
DB 1848 CCGCTCTCGCTCCCAAGTGTCTGGATTAAGGTTGTAGGCCACCTCGCTGGCGCTCTT 1789
QY 1693 TTCTCTTTTGTGAGTAGGGCCCCCAGACCAAAATAATAATCTTAGAATCAAAATCAG 1752
DB 1788 GGTCTTTTCTTTCAAAACACGTGGGCCAGCCAGATATCTTTATTTTCGCGAAAGGAACCTC 1729
QY 1753 TGTGTGTGTTTG-ACCACCTGTCACTTGAGAACCAACAGTGTGACAGGCGCTCAGAGT-- 1809
DB 1728 GGTGAGAAATAGCAAGACAGTTACTCAAGATCATGTGATATAACCGCTAGTTGTAATTC 1669
QY 1810 --AGAGGTGATCTGTCTCGAAAGAGAAATAGATGAATAATTTCTCGGGCCAGCGTG 1867
DB 1668 CAACATTTGSAATCTGACCTCCAGGATGSCACCTGTTAGAAAGAGTGGATTTGGGCCAGCGTG 1609
QY 1868 --GTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCACTGAG 1925
DB 1608 CTGTGGCTCAACCTGTAAATCTAGCGCTTTGGGAGGACAGGCGAGCTGATTTGCTTGA 1549
QY 1926 GTGAGAGTTCAAAACCCAGCTGGCCAAACATGTTGAAAACCCCGTCTCTACTAAAAATACA 1985
DB 1548 GTGAGTAGTTAATAAATAGCTGTGTCACATGTTGAAAACCTCATCTACTAAAAATACA 1489
QY 1986 AAAAATTAGCTAAGTGTGTGGCGCATGCCGTAAATCCAGCTACTTGGAGGGTGGAGC 2045
DB 1488 AAAATTAGCC--AGCGGTGTGGCAACCCATATAGTCCAGCTACTTGGGGGCTGAGGC 1431
QY 2046 AGGGAATTTCTTGAACCCGGGAGGAGAGAGTGTGAGTGAAGCGAGATCACACCACTGCA 2105
DB 1430 AGGGAATCTGCTTGAACCCGGGAGGTTGGAGGTTGAGGAGGCCAGATC-TGCCACTGCA 1372
QY 2106 CTCAGCTGGGGGAGAGAGCGAGATCTTCTCTCAAAAAACAAAAACAAAGAAAT 2162
DB 1371 CTCTAGCCTGAGGAGACAGAGCAAGATCTCCTCTCAAAAAACAAAAAGAAAAAGT 1315
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 281.273 Seconds
(without alignments)
7714.406 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910

Perfect score: 3910

Sequence: 1 aggaattcattccatttaaat.....cggacacggtcacccgttgc 3910

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgm2_6/ptodata/2/ina/5A.COMB.seq.*

2: /cgm2_6/ptodata/2/ina/5B.COMB.seq.*

3: /cgm2_6/ptodata/2/ina/6A.COMB.seq.*

4: /cgm2_6/ptodata/2/ina/6B.COMB.seq.*

5: /cgm2_6/ptodata/2/ina/PCFUS.COMB.seq.*

6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349.8	8.9	12394	4	US-09-488-856A-10
C	345	8.8	84495	4	US-09-797-906-3
C	322.4	8.2	3805	4	US-09-108-006C-3
4	321	8.2	35060	3	US-08-814-095-7
5	319.8	8.2	392000	4	US-10-027-983-11
C	316.6	8.1	59065	4	US-09-813-817-3
C	316.6	8.1	59065	4	US-09-978-197-3
8	312.4	8.0	174493	4	US-09-804-471A-3
9	312.4	8.0	174493	4	US-10-238-709-3
10	308.2	7.9	162450	4	US-09-345-882-1
11	307.6	7.9	43950	4	US-09-735-934A-3
12	307.6	7.9	43950	4	US-10-060-332-3
C	307.2	7.9	70000	4	US-09-851-896-3
C	304	7.8	39982	4	US-09-820-924-3
C	303	7.7	43950	4	US-09-733-934A-3
C	303	7.7	43950	4	US-10-060-332-3
17	301.8	7.7	246240	2	US-08-724-394A-20
18	301.8	7.7	246240	2	US-08-724-394A-21
19	301.8	7.7	246240	2	US-08-724-394A-22
20	299.8	7.7	39982	4	US-09-820-924-3
21	297.2	7.6	70000	4	US-09-851-896-3
22	296.6	7.6	59065	4	US-09-813-817-3
23	296.6	7.6	59065	4	US-09-978-197-3
C	291.2	7.4	83450	4	US-09-811-469-3
C	289.2	7.4	55298	4	US-09-491-356C-1
26	288	7.4	21234	4	US-09-810-671-3
27	288	7.4	21234	4	US-10-109-854-3

28	287.4	7.4	116592	4	US-09-818-512-3	Sequence 3, Appli
C	286.8	7.3	35060	3	US-08-814-095-7	Sequence 7, Appli
30	285.8	7.3	162450	4	US-09-345-882-1	Sequence 1, Appli
C	283.8	7.3	12394	4	US-09-488-856A-10	Sequence 10, Appli
32	282.8	7.2	3805	4	US-09-108-006C-3	Sequence 3, Appli
C	282.8	7.2	246240	2	US-08-724-394A-20	Sequence 20, Appli
C	282.8	7.2	246240	2	US-08-724-394A-21	Sequence 21, Appli
C	282.8	7.2	246240	2	US-08-724-394A-22	Sequence 22, Appli
C	280	7.2	116592	4	US-09-818-512-3	Sequence 3, Appli
37	275	7.0	83450	4	US-09-811-469-3	Sequence 3, Appli
38	275	6.9	84495	4	US-09-797-906-3	Sequence 7, Appli
39	270.4	6.9	72604	4	US-09-268-992-7	Sequence 7, Appli
40	270.4	6.9	72604	4	US-09-857-474-7	Sequence 1, Appli
C	269.4	6.9	23187	4	US-09-499-522-1	Sequence 6, Appli
42	267.2	6.8	45546	4	US-09-146-053-6	Sequence 41, Appli
C	265.6	6.8	21721	4	US-09-269-939A-41	Sequence 42, Appli
44	264	6.8	44848	4	US-09-435-739-42	Sequence 10, Appli
C	263	6.7	98844	4	US-09-791-211-10	

ALIGNMENTS

RESULT 1

US-09-488-856A-10
; Sequence 10, Application US/09488856A
; Patent No. 6316259

GENERAL INFORMATION:

APPLICANT: Brett P. Monia
APPLICANT: Robert McKay
APPLICANT: Madeline M. Butler
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP
FILE REFERENCE: RTS-0115
CURRENT APPLICATION NUMBER: US/09/488.856A
CURRENT FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 88

SEQ ID NO 10

LENGTH: 12394

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (115)...(397)

NAME/KEY: CDS

LOCATION: (2438)...(2625)

NAME/KEY: CDS

LOCATION: (5639)...(5722)

NAME/KEY: CDS

LOCATION: (5864)...(5974)

NAME/KEY: CDS

LOCATION: (7902)...(8032)

NAME/KEY: CDS

LOCATION: (8121)...(8227)

NAME/KEY: CDS

LOCATION: (9197)...(9294)

NAME/KEY: CDS

LOCATION: (9375)...(9470)

NAME/KEY: CDS

LOCATION: (9898)...(10084)

NAME/KEY: CDS

LOCATION: (10431)...(10523)

NAME/KEY: CDS

LOCATION: (11713)...(11786)

US-09-488-856A-10

Query Match 8.9%; Score 349.8; DB 4; Length 12394;

Best Local Similarity 55.2%; Pred. No. 5.9e-77;

Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;

QY 370 TATTATTAGACAGAGTTCTCACTGTGACCCAGGCTGGAGTGCAGTGGCACAATCATTA 429

DB 3604 TTTTITTTGAGATGAGTCTTACTCTGTGACCCAGGCTGGAGTGCAGTGGCAGATCTTG 3663

Query Match	8.8%	Score 345	DB 4	Length 84495	
Best Local Similarity	55.0%	Pred. No. 2.6e-75			
Matches 979	Conservative	0	Mismatches 715	Indels 85	Gaps 12
445	AAACTCTGGGCTCAAGTGATCCTACTACTCAGCC-TCCAGAGTAGCTAGGACTACAGG	503			
18461	AACTTCTGGGCTGAAGCCATCCACTGCGCTTGGCGCTCCCAAGTGGTGGGATACAG	18402			
504	CAACACAGCCATACCTGCTAAATTTTTTTTTTAAATTTTCAATTTATGTATTCATTTTC	563			
18401	CATTAGTCCAGCCCTGGACTATTTATATATATATATATATATATATATATATATATTTT	18342			
564	TTTCTTTTGTG-TTGTGTGTGTTTGAGATAGGGTCTCATTTGTTTACCCAGGCTGGAGG	622			
18341	TTTTTTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCATCTCTGTCAACCAGGCTGGAGT	18282			
623	GCAGTGGCATGGTGA CAGCTGA--GCAGCTTTGACTTCTCTGGGCTCAAGTGATCTCCCTG	680			
18281	GCAGTGGCATGATCTCAGCTCACTGCAACCTTGGACTCTCTGGGTTCAAGCAATTTCTCTG	18222			
681	CCTCAGGCTCCCAAGTAGCTGGGACTACAAACACGTGTCAACATGCTCGCTGGCTGATATTTT	740			
18221	CCTCAGGCTCTTAAGTASCCTGGGATTTACAGGCACATGTGACCGCGCCAGCTAAATTTTTT	18162			
741	TTTTCTTTAAACAGGGTATCATCTGTGTGGCCAGGCTGGAGTACAGTGGCGTAAATAATAG	800			
18161	TTTTCTT-----TTTTTGTTTTTTTTTTTTGTATAGACGGAGTTTCCTATATATTAG	18109			
801	CTCACTGCAGCTCCCTCTGGGCTCAAGCAATCCGCTGGGCTCAGCATCTCTGAGTAGC	860			
18108	CCAGGCGGATCTGAACTCTCTGACCTCAGTGATCTGCGCGCTCAGCCTCCCAAAGTGT	18049			
861	TGGGACTACAGGCTGTGTGCCACAGGCCAGCTAAGTTTAAAAAATGATTTTGGTATA	920			
18048	TAGATTACAGGCATGAGCCACCAATGCGCGGCTCTCTATTTTATAGTTTAAAGACAA	17989			
921	GAGAGGCTTTGCTATGTGTCTCAGGCTGTATTTTATTTGTAGACAGGCTCTCACTAT	980			
17988	GA----GTCCCACTCTGTGTGCCAGGCTGGAGTGCAGTGGCACAAATCATAGTCACTGCA	17933			
981	GTGCGCATGATCCCCCACTCCACTTCCCAAGTGCTCATCTTATCTG--TTCAATTAGT	1038			
17932	GCTTTAAACTCTGGGCTAAAGCAATCCTACCTCAGCCTCCGAGTAGCTACATGCAC	17873			
1039	CAGTTGACAGACATTTAGGTTGTTTCCACTTTTTTGACCAATATGAATAAATCTCCAGTGA	1098			
17872	ATGCCACCATACCCAGCTAAATTTTTTAAATGTTTTTAAAGATGGGGTCTCACTATGTTG	17813			
1099	ATATTCATGTATACATTTGTGTGGGCATATGTTTTCATTTCTGTGGGTTTATATCTAGG	1158			
17812	CCCAGGCTGCTCGAACTCTCTGGGCTTTAAGCGATCTCTCCCATCTGTCGTTTCCAAAGTG	17753			
1159	AGTGGAAATTCGTGGATCCCGGTATATTTTGACAGGCAGAGTTTCAGGGGGAAGAAAAC	1218			
17752	CTGGGATTAACAAGTGTGAGCCACCATGTCCTGCCCAATGTACTTTTTTCAAGGTCAC	17693			
1219	TGGGAAATTAAGACATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTTCGAGATTATTT	1278			
17692	GACTGCAGTGTGG--AGATAAATCGGAGGAAGTATAATAGTTTAAATAGGCAGTATTTTA	17635			
1279	TTATATCTGTGTGACAAATGTGCGAGTTTGAATGAAGNATCAAGTATATCTAAGTGAGAAGT	1338			
17634	ATGACCATACACAGGCCATCTCAACTATATGAATAAATCATTAATTTACTGATCATATTTCT	17575			
1339	GAGAAATTAAGGCTGGAATAGGGCGTTTCAGAGTAAAAATCATGAAGCACTTTTGA--	1394			
17574	CTATTTTATTTGGACATTTAGACTATTTCCAGTTCTTTTGCTAATATTAATATGAAGGACAC	17515			
1395	CAAAAATTAAGGAGCTTTGGCTGTAAACAAATAATAAAAAATCAAAATTTTTTTTTTTTTT	1454			
17514	TGTTAATAAACATTTCTTGAACCTAAATCTTTGGCAGCATCTCTAAATTAATTTTCTTTCTTT	17455			
1455	TTGAGAAAGAGTCTTGCTCTTTTCACTCTGGCTGGGCGAGTGTGTGATCTCAGCTCAC	1514			

RESULT 3
US-09-108-006C-3/c
Sequence 3, Application US/09108006C
Patent No. 6524613
GENERAL INFORMATION:
APPLICANT: Steer, Cliford J.
Kren, Betsy T.
Bandhopadhyay, Paramita
Roy-Chowdhury, Jayanta
Chimeraplasy
TITLE OF INVENTION: Hepatocellular
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kimeragen, Inc.
STREET: 300 Pheasant Run
CITY: Newtown
STATE: PA
COUNTRY: USA
ZIP: 18940
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/108,006C
;   FILING DATE: 30-Jun-1992
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/054,288
;   FILING DATE: 30-APR-1997
;   APPLICATION NUMBER: 60/054,837
;   FILING DATE: 05-AUG-1997
;   APPLICATION NUMBER: 60/064,996
;   FILING DATE: 10-NOV-1997
;   APPLICATION NUMBER: 60/074,497
;   FILING DATE: 12-FEB-1998
;   APPLICATION NUMBER: PCT US 98/08834
;   FILING DATE: 30-APR-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Friebel, Thomas
;   REGISTRATION NUMBER: 29258
;   REFERENCE/DOCKET NUMBER: 7991-015-999
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 215-504-4444
;   TELEFAX: 215-504-4545
;   TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 3805 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: Genomic DNA
;     FEATURE:
;       NAME/KEY: exon
;       LOCATION: 71...114
;       OTHER INFORMATION: Exon 1
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-108-006C-3

```

```

Query Match      8.2%; Score 322.4; DB 4; Length 3805;
Best Local Similarity 68.3%; Pred. No. 2.2e-70;
Matches 496; Conservative 0; Mismatches 216; Indels 14; Gaps 3;

QY 1446 TTTTGTGAGAAAGAGTCTTCCCTTCCAGCAATTCCTGCTTCCAGCTCCCAA 1505
DB 1876 TATGATATTTAAACAGGGTCTTCTCTGTCACCCAGGCTGGAGTGGCGGCTGAAC 1817
QY 1506 TCAGCTCACTGCACTTTCGGCTCCCGGGTTCAAGCAATTCCTGCTTCCAGCTCCCAA 1565
DB 1816 ATGGTTCACCTGCACTTTCGGCTCCCGGGTTCAAGCAATTCCTGCTTCCAGCTCCCAA 1575
QY 1566 GTAGCTGGGACTACAGGCACTTCCACCAGTCCAGCTGATTTTGTATTTTAGTAGAG 1625
DB 1756 GTAGCTGAGAGCACAGGTTGGGCAACCATGCTGCTAAATTTTGTATTTTAGTAGAG 1697
QY 1626 ATGGATTTCACTTTGTTGGCCAAAGCTGGTCTCAAACT-----TTTGTCTGTC 1673
DB 1696 ACAGGGTCTCACTATGTTGCCAGGCTGGTGTGAATCTCCTGGCTCAAGTGATCCTCCC 1637
QY 1674 ATAAATTTGTAACATTTATTTCTTTTCTGAGGTAGGCCCCCAGACCAAAAAATAA 1733
DB 1636 ACCTTGGCTCCCAAGTCTGGGATTAACAGGGGTGAGCCAGCCGCCAGCGGAAAGCA 1577
QY 1734 ATCTTAGAATCCAAATCAGTGTGTTGGTTGACCACTGTCACTTGAGAACACAGTGTGA 1793
DB 1576 TGTATTTAGTGTCTTGTGTGCCCCAGGAGGCTGTGTGGGGGTGATGGAGATAAGA 1517
QY 1794 CAGGGCTCAGGAGTGTAGGTTGATCTCTGTCGAAAGAGAAATAGAAATGAAATAT-TC 1852
DB 1516 TCACAGCTGCCCCGGTGTCTGTTGTTTCACTATCTGCTGCAATGCAATTAGAAACCTTAA 1457
QY 1853 TCCGGGCGAGGGGTGGTCTCATCTGTTATCCAGCACTTTTGGGAGGCCAAGGCATG 1912

```

```

DB 1456 TCCAGCCAGGTGCGGTGGCTCACGCTGTAAATCCAGCACTTTTGGAGATCGAAACGGG 1397
QY 1913 TGGATCACCTGAGGTGAGGATTTCAAAACAGCTGGCCCAACATGTTGAACCCCGTCTC 1972
DB 1396 CAGATCACCTGAGGTGAGGATTTCAAGCCAGCTGGCCCAACATGTTGAACCCCGTCTC 1337
QY 1973 TACTAAAAATACAAAAA--TTAGCTAAGTGTGGTGGCGCATCCCTGTATATCCAGCTACT 2031
DB 1336 TACTGAAAAATACAAAAAAGTTAGTCGGGTGTGGTGGCATGTGCTGTAGTCCAGCTACT 1277
QY 2032 TGGAGGGGTGAGGACAGGAGATTTCTTGAACCCCGGAGGACAGGTTGCAAGTGAAGCGAG 2091
DB 1276 TGGGAGGCTGAGGACAGGAGATTTGGACCTTGGGAGGCGGAGCTTGCAGTGAGCGGAG 1217
QY 2092 ATCAACACCACTGCACTCCAGCCTGGGAGGAGAGAGAGACTTCTCTCAAAAAACAAAA 2151
DB 1216 ATCCGCGCACTGCACTCCAGCCTGGGAGGACAGAGGAGACTTCTCTCAAAAAACAAAA 1157
QY 2152 AACAA 2157
DB 1156 CAACAA 1151

RESULT 4
US-08-814-095-7
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
;   APPLICANT: Sored, Hermona
;   APPLICANT: Zakut, Haim
;   APPLICANT: Shani, Moshe
;   TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;   TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: KOHN & ASSOCIATES
;   STREET: 30500 No. 6025183thwestern Highway, Suite 410
;   CITY: Farmington Hills
;   STATE: Michigan
;   COUNTRY: U.S.
;   ZIP: 48334
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/814,095
;   FILING DATE:
;   CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Montgomery, Ilene N.
;   REGISTRATION NUMBER: 38,972
;   REFERENCE/DOCKET NUMBER: 2391.00066
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (248) 539-5050
;   TELEFAX: (248) 539-5055
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 35060 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: other nucleic acid
;     DESCRIPTION: /desc = "Cosmid including ACH
;     DESCRIPTION: promotor, ACH gene and ARS gene"
;     HYPOTHETICAL: NO
;     ANTI-SENSE: NO
;     ORGANISM: Homo sapiens
;     POSITION IN GENOME:
;     CHROMOSOME/SEGMENT: 7q22
;     FEATURE:

```

```

NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: 34528..34895
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: Gene"
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 34092..34358
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 33779..33963
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 33493..33591

```

```

OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 33297..33408
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: 32959..33094
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: 32569..32628
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: 32386..32468
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: 31894..32080
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: 31363..31534
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: 30816..31011
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: 30470..30626
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: 30187..30274
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: 29945..30073
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: 29664..29856
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7

```

```

Query Match      8.2%; Score 321; DB 3; Length 35060;
Best Local Similarity 68.8%; Pred. No. 1.6e-69;
Matches 511; Conservative 0; Mismatches 205; Indels 27; Gaps 4;
QY 1445 TTTTTCCTTCTTGAAGAGTCTTCTTTCACCTGGCTGGAGGCGCATGGTGTGAT 1504
      |||||
Db 9802 TCTTTTCTTTTGAGACGGAGTCTCGCTCTATCGCCAGGCTGGAGTGCAGTGGCAAT 9861
      |||||
QY 1505 CTCAGCTCACTGCAACTTTCGCCCTCCCGGGTTCAAGCAATTCCTCTGCTTCAGCCCTCCA 1564

```

```

Db      9862 CTCGGCTCACTGCAACCTCTGCGCACCGAGTTCAAGCAATTTCTCTGTGTGTCAGCCCTCCCA 9921
QY      1565 AGTAGCTGGAGACTACAGGCACTTCCACCATGCGGAGCTGATTTTGTATTTTGTAGTAGA 1624
Db      9922 AATAGCTGGGATTACCAAGTCCCAACCAATATGCCAGCTAAATTTTATATATTTTCAAGTAGA 9981
QY      1625 GATGGGATTTCACTTTGTTGGCAAGCTGCTCTCAAACTTTTGTGCTGTCTATAATTTGTGT 1684
Db      9982 GACAGGTTTCAACATGTTGGCCAGGCTGTCTCAAAATCTCTGCTGCTCAAGTGTATCTGCC 10041
QY      1685 AAC-----TATTTCTCTTTGCTGTAGGTTAGGSCCCCGCCAGACCAAAAAATA 1732
Db      10042 CACTCAGCCTCCCAAGTGTGGATTACAGATGTGAGCAGCTCACTCAGCCAGAGATC 10101
QY      1733 AATCTTAGATCCAAATCACTGTGTTGGTTTGACCACTGTCACTTGAGAACCA----- 1785
Db      10102 CCTTCCAAATTAATAAATCAACAGAAACCCCAAGAAATCAATTTGTTGGCAATGCTCT 10161
QY      1786 -----CAGTGTGACCAAGGCGCTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAG 1839
Db      10162 AGTAATATGTGCATCTTGGGTGCAAGCGTCTAAATTTCTAAACCTGCATCTGCAGTTTAA 10221
QY      1840 AATGAATAATTTCTCGGGCAGCGGTGGTCTCATGCTCTAATCCAGCAGCTTTGGG 1899
Db      10222 CATTTTAAGATGCTTCCGGCCAGCGGTGGTCTCAGCCTGTAAATCCAGCAGCTTTGGG 10281
QY      1900 AGCCAGGAGCATGTGATCACTCAGGTGAGGAGTTCAAACCAAGCTGCGCCCAACATGTT 1959
Db      10282 AAGCCGAGGAGGTGGATCACTTCCAGTCCAGGAGTTCAGACTAGCTGACCAACATAGT 10341
QY      1960 GAAACCCCGCTCTCTACTAAAAATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTGTA 2019
Db      10342 GAAACCCCACTCTACTAAAAA-ATACAAAAATTAGCCAGGTGTGGTGGCGCATGCTCTGTA 10400
QY      2020 ATCCAGCTACTTGGGAGGTGGAGGAGAGAAATTTCTGAACCCGGAGGAGGAGGTTG 2079
Db      10401 ATCCAGCTACTCAGGAGGCTGAGGAGGAGAAATTTGTTGAACCCGGAGGAGGAGATTG 10460
QY      2080 CAGTAGAGCAGATCACACCACTCACTCCAGCCTGGGGGAGA-GAGCGAGACTTCTCT 2138
Db      10461 TAGTGACTGAGATTGCACCATGTCATCTCCATCTCTGGCAACAGGAGCAAACTCCGCT 10520
QY      2139 CAAAAAACAACAAAAACAAAAAGAA 2161
Db      10521 CAAACAACAACAACAACAGAAA 10543

RESULT 5
US-10-027-983-11
; Sequence 11, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137740
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 137742
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (138122)...(138221)
; OTHER INFORMATION: exon 10:intron 10

;
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 145507
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 151967
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (151967)...(1542066)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 154217
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (164037)...(164136)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (174657)...(174756)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (186224)...(186323)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (195242)...(195341)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 202703
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (202771)...(202870)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (206246)...(215602)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (218126)...(218225)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (220360)...(220459)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (222717)...(222816)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (223981)...(224080)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (227487)...(227586)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (230157)...(230256)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (232299)...(232398)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (236552)...(2366651)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (238789)...(248788)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: exon
; LOCATION: (118288)...(119101)
; OTHER INFORMATION: exon 1C
; NAME/KEY: exon:intron junction
; LOCATION: (151129)...(151130)
; OTHER INFORMATION: exon 5:intron 5
; NAME/KEY: exon:intron junction
; LOCATION: (299248)...(299249)
; OTHER INFORMATION: exon 9:intron 9
; NAME/KEY: exon:intron junction
; LOCATION: (348578)...(348579)
; OTHER INFORMATION: exon 10:intron 10

```

```

; NAME/KEY: intron
; LOCATION: (348579)....(381838)
; OTHER INFORMATION: intron 10
; NAME/KEY: intron:exon junction
; LOCATION: (386185)....(386186)
; OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Query Match      8.2%; Score 319.8; DB 4; Length 392000;
Best Local Similarity 67.2%; Pred. No. 1.1e-68;
Matches 533; Conservative 0; Mismatches 217; Indels 43; Gaps 4;

QY 1419 ACAAAATAATAAAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGAGCTTCTCTTTCA 1478
Db 66965 ACTAAATTTTTTTCCTATCTCACATTTCTTTCTTTTGGAGACAGAGTATTACTCTGTCA 67024

QY 1479 CCTTGGCTGGAGGCGAGTGGTGTGATCTCAGCTCACTGCAACTTTTCGGCTCCCGGGTTCA 1538
Db 67025 CTTAGGCTGGAGGCGCAGTGGCGCAATCTTGGCTCACTGCAACTCTCTGGGTTC 67084

QY 1539 AGCAATTTCTCTGCTTTTCAGCTCCCAAGTAGCTGGGACTACAGGCACTTCCACCATGGC 1598
Db 67085 AGCAATTTCTCTGCTCGGCTCCCGAGTAGCTGGGATTAGAGGCGGCTGCCACCATGCC 67144

QY 1599 CAGCTGAATTTGTATTTTATAGAGATGGGATTTCACTTTTGGGCCAAGCTGGTCTC 1658
Db 67145 TGGCTAAATTTTGTATTTTATAGTAGAGCAGGGTTTCATCAGTTGGTGGGCTGGTCTC 67204

QY 1659 AACT-----TTTTCCTGTCAATTTGTGTAATTTGTAATTTGTTCTTTTCTGAG----- 1706
Db 67205 CAACTCCTGACCTTGGCTTCCCAAGGTCTGGTATTTACTTTCTAACCTCTGTGTATGCT 67264

QY 1707 -----GTAGGCGCCCCAGACACCAAAAAATAAATCTTAGAATCCAAATCAGTG 1754
Db 67265 ACCTATTATTATTTGTCTCTCTACACTAGATAATAAGCTCTATGAGGGTAGACTTCTT 67324

QY 1755 TGTGTGTTGACCACTGTCTACTTGAGAACACAGTGTGACAGGGCTCAGAGTAGAGG 1814
Db 67325 TGTTTGTCCAATGCTCTATTTTCCAATATTTATAACCGTACTGGCCGGTAGGTACTGTCA 67384

QY 1815 TGAATCTGTCTCAAGAG-----AAATAGAATGAAATATTTCTCG 1856
Db 67385 ATTATAGTTTTTGAATAAATTTAGAGTAAACAAATTTAACTTGTGAAATAAATAATG 67444

QY 1857 GGCAGCGGTGGTGGCTCATGCCCTGTAAATCCAGCACTTTGGGAGGCCAAGGCATGTGA 1916
Db 67445 GGCAGCGGTGGTGGCTCAGCCCTGTAAATCCAGCACTTTGGGAGGCCGAGGCGGCAGA 67504

QY 1917 TCACCTGAGCTCAGGAGTTCAAAACACAGCCTGCCAACATGFTGAAACCCCGTCTCTACT 1976
Db 67505 TCACCTGAGCTCAGGAGTTCAAGACCAAGCCTGGGCAACATGFTGAAACCCCGTCTCTATT 67564

QY 1977 AAAAATACAAAAATTAGCTTAAGTGTGGTGGCGCATGCCGTGTAATCCAGCTACTTTGGGA 2036
Db 67565 AAAAATAC-AAAAAATAGCTGGGTGGTGGTGCATGCCCTGTAATCCAGCTACTCGGA 67623

QY 2037 GGGTGGGCGAGGAGAAATTTCTTGAACCCGGAGGCGAGAGGTTGCACTGAAGCGAGATCAC 2096
Db 67624 GGTGAGGCAACAAGATCGCTTGAACCTTGAGGCGAAGGTTGCACTGAGCGCGCATGCA 67683

QY 2097 ACCACTGCATCCAGCCTGGGGGAGAGCGAGACTTCTCTCTCAAAAAACAAAAACAA 2156
Db 67684 GCCACTGCATCCAGCCTGGGCGAGCAGCGAGACTCTGTCTCAAAATAAATAAATAA 67743

QY 2157 AAGAAATTAAGCAA 2169
Db 67744 TAAAAATTAATAA 67756

```

RESULT 6
US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583

QY 1801 CTCAGGAGTACAGGAGTCTCTGCTCGAAGAGAAATAGAAATATTCTCCGGGCC 1860
DB 126735 GTCCTTCGAACCTTGTTATTTCTCCCTGTTAATACAGTAAATTAACTAAGTTGGCC 126794
QY 1861 AGGCGTGTGCTCATGCTGTAATCCAGACACTTTGGAGGCCAAGGCATGTGGATCAC 1920
DB 126795 AGGAGTGTGGCACATGCTGTAATCCAGACACTTTGGAGGCCAAGGCATGTGGATCAC 126854
QY 1921 CTGAGGTACAGGAGTTCAAAAACAGGCTGGCCAAACATGCTGAAACCCCGCTCTCTACTAAAA 1980
DB 126855 CTGAGGTACAGGAGTTGAGACAGGCTGGCCAAACATGCTGAAACCCCGCTCTCTACTAAAA 126914
QY 1981 ATACA----AAAAATAGCTAAGTGTGTGGGCGCATGCTGTAATCCAGACTTTGGGA 2036
DB 126915 AAAAATTTGAAAAAGTTAGCTGGGTGTGGTGTGGCGCACCTGTAGTCCAGACTATTCCGGA 126974
QY 2037 GGGTACGAGGAGAAATTTCTTGAACCCGGAGGAGGAGGTTGCAAGTGAAGCGATCAC 2096
DB 126975 GGCCTGAGGAGGAGAAATTTGCTTGAACCCAGGGGCTGGAGGTTGCAAGTGAAGCGATCAC 127034
QY 2097 ACCACTGCACCTCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2156
DB 127035 GGCACCTGCACCTAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 127093
QY 2157 AAGAATTAAGCAAAATAGACATTTGAGAGAGAACTTGAAGGGGTCAGACCAAGTACAGA 2216
DB 127094 AGTTAAGTAGTACTTTGGGGCCCTATCAGATAGTGTCTCTCGAGGAGGCGCTGGGCCCTTA 127153
QY 2217 TTTCTGTGCCACAT 2230
DB 127154 CTACAGTTTCACTT 127167

RESULT 9

US-10-238-709-3
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match 8.0%; Score 312.4; DB 4; Length 174493;
Best Local Similarity 61.7%; Pred. No. 5.2e-67;
Matches 490; Conservative 0; Mismatches 298; Indels 6; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGTGAGAAAGCTCTGCTCTTTCACCGCTGGCTGGAGGCGAGTGGTG 1500
DB 126376 TCTCTCTTTCTTTTCAAGACAGAGTCTTGTCTGTGACCCAGGCTGGAGTGGAGTGG-G 126434
QY 1501 TGAATCTCAGTCACTCAACTTTGCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAAGCT 1560
DB 126435 CGACCTTGGCTCTACTGCAACCTCCGGCTCTTGGGTTCAAGAGATTTCTCTGCTCAGCTC 126494
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGCCCGAGTGAATTTTGTATTTTAG 1620
DB 126495 CCCGAGTAGCTGGGATTACAGGTCGCCGCCCATGCCTGGCTAAATTTTGTATTTTAG 126554

QY 1621 TAGAGATGGGATTTTCACTTTTGTGCGCCAGAGCTGGTCTCAAACTTTTTCGTGTCATAATTG 1680
DB 126555 TAGAGATGGGATTTTCAACCATGTTGAAGCATTTTATAGAAAGGGGTTAGGAGAGGAGACA 126614
QY 1681 TTGTAACATTATTTGTTCTTTTCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
DB 126615 NNN 126674
QY 1741 AATCCAAATCAGTGTGTTGGTTTGACCACTGTCTACCTTGAGAACCAAGTGTGACAGGGC 1800
DB 126675 NNTCTCCCTTTCTCCCAAAATATGCGACACTCTTCTGTCCCCCTAGTCAATTCTCATATA 126734
QY 1801 CTCAGGAGTACAGGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGGGCC 1860
DB 126735 GTCCTTCTGAACTTGTGTTTATTTCTTCTCCCTGTAATACAGTAAAAATTAATTAAGTTGGCC 126794
QY 1861 AGGCGTGTGCTCATGCTGCTGTAATCCAGACACTTTTGGGAGGCCAAGGCATGTGGATCAC 1920
DB 126795 AGGAGTGTGGCACATGCTGTAATCCAGACACTTTTGGAGGCCGAGGCGAGCAGATCAC 126854
QY 1921 CTGAGGTACAGGAGTTCAAAAACAGGCTGGCCAAACATGTGAAACCCCTCTCTACTAAAA 1980
DB 126855 CTGAGGTACAGGAGTTTGAACAGGCTGGCCAAACATGTGAAACCCCTCTCTACTAAAA 126914
QY 1981 ATACA----AAAAATAGCTAAGTGTGTGGCGCATGCTGTAATCCAGACTTTGGGA 2036
DB 126915 AAAAATTTGAAAAAGTTAGCTGGGTGTGGTGTGGCGCACCTGTAGTCCAGACTATTCCGGA 126974
QY 2037 GGGTACGAGGAGAAATTTTGTAAACCCGGAGGAGGAGTGTGAGTGAAGCGAGATCAC 2096
DB 126975 GGCCTGAGGAGGAGAAATTTGTAACCCAGGGGCTGGAGGTTGCAAGTGAAGCGATCAC 127034
QY 2097 ACCACTGCACCTCAGCTGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2156
DB 127035 GGCACCTGCACCTAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 127093
QY 2157 AAGAATTAAGCAAAATAGACATTTGAGAGAGAACTTGAAGGGGTCAGACCAAGTACAGA 2216
DB 127094 AGTTAAGTAGTACTTTGGGGCCCTATCAGATAGTGTCTCTCGAGGAGGCGCTGGGCCCTTA 127153
QY 2217 TTTCTGTGCCACAT 2230
DB 127154 CTACAGTTTCACTT 127167

RESULT 10

US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 639373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele

```
,
, LOCATION: 88073
, OTHER INFORMATION: 5-127-261 : polymorphic base A or C
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90842
, OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93714
, OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97122
, OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97152
, OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99098
, OTHER INFORMATION: 5-130-257 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99117
, OTHER INFORMATION: 5-130-276 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 103806
, OTHER INFORMATION: 5-131-395 : polymorphic base A or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 106940
, OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108106
, OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108149
, OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108308
, OTHER INFORMATION: 5-135-357 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 108471
, OTHER INFORMATION: 5-136-174 : polymorphic base C or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134134
, OTHER INFORMATION: 5-140-120 : polymorphic base C or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134362
, OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
, FEATURE:
, NAME/KEY: allele
, LOCATION: 134374
, OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
, FEATURE:
, NAME/KEY: allele
, LOCATION: 146328
, OTHER INFORMATION: 5-143-84 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 146345
, OTHER INFORMATION: 5-143-101 : polymorphic base A or C
, FEATURE:
, NAME/KEY: allele
, LOCATION: 150329
,
,
, OTHER INFORMATION: 5-145-24 : polymorphic base A or G
, FEATURE:
, NAME/KEY: allele
, LOCATION: 160031
, OTHER INFORMATION: 5-148-352 : polymorphic base G or T
, FEATURE:
, NAME/KEY: allele
, LOCATION: 72771..72817
, OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
, FEATURE:
, NAME/KEY: allele
, LOCATION: 72771..72817
, OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
, FEATURE:
, NAME/KEY: allele
, LOCATION: 88050..88096
, OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
, FEATURE:
, NAME/KEY: allele
, LOCATION: 88050..88096
, OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90819..90865
, OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
, FEATURE:
, NAME/KEY: allele
, LOCATION: 90819..90865
, OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93690..93736
, OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
, FEATURE:
, NAME/KEY: allele
, LOCATION: 93690..93736
, OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97099..97145
, OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97099..97145
, OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97130..97177
, OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
, FEATURE:
, NAME/KEY: allele
, LOCATION: 97130..97177
, OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99075..99121
, OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99075..99121
, OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99094..99140
, OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
, FEATURE:
, NAME/KEY: allele
, LOCATION: 99094..99140
, OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
, FEATURE:
, NAME/KEY: allele
, LOCATION: 103783..103828
, OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
, FEATURE:
```

FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match 7.9%; Score 308.2; DB 4; Length 162450;
Best Local Similarity 66.7%; Pred. No. 5.6e-66;
Matches 529; Conservative 0; Mismatches 239; Indels 31; Gaps 5;
QY 1421 AAAATAATAAAAAACACAAATTTTTTTTTTTTTTTTGTGAGAGAGCTGTGCTCTTTTACC 1480
DB AATGTAAACAAAGTGAACAACTATTTTTTCTTTTGTAGGGGGAGCTAGCTGT-TGC 86793
QY 1481 CTGGCTGAGGGAGTGTGTGATCTCAGCTCAGTGCACACTTTCCCTCCCGGTTCAAG 1540
DB CGGGCTGAGTGCAGTGGCACCACCTTGGCTCAGTGCACAGCTCCGCCCTCCTGGGTCAAG 86733
QY 1541 CAATCTCTCTCAGCTCCAGCTCCAGTACAGTGGGACTCAGCACTTCCACCATGCCCA 1600
DB CAATCTCTCTCAGCTCCAGCTCCAGTACAGTGGGACTCAGCACTTCCACCATGCCCA 86673
QY 1601 GCTGATTTTTGTATTTTAGTAGAGATGGGATTTCACTTTGTGCCCCAGCTGCTCAA 1660
DB GCTAATTTTTGTATTTTGGTAGAGAGAGGGTTTCCACCATGTTAGCCAGGATAGTCTCGA 86613
QY 1661 ACT-----TTTTGCTGTCATAATTTGTTGTAACATTTGTTTCTTTTGTGCTGAGT 1708
DB TCTCTGACCTGCTGATCCGCCCACTCTCGCTCCCAAAAGTGTGGGATTTAGGGCTGAG 86553
QY 1709 AGGGCCCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGTTTGACCA 1768
DB CCACCCCACTGGCCAAAAAAGTAAAACTATATAATAATTATTAAGCTTTTAAAGTGA 86493
QY 1769 CTGTCACCTTGAACACACAGTG-----TGACAGGGGCTCAGAGTAGAGTGAATCTC 1821
DB CATGGGAATACCAACCCCTTGAATTTCAATGATCAGTCAATAAATTTAAATAATTTGCTC 86433
QY 1822 TGCTGAAGAGAAATAGATGAATAATTTCTCCGGGCCAGG-----CGTGGTGG 1871
DB TGCGAGAAACAAAGCAAAATCTTTCAAGCTTGTCTATTGAATGAAGTACTGTACTTTGGCT 86373
QY 1872 CTCATGCTGTAAATCCAGCACTTTGGGAGGCCAAGGATGGGATCACTCAGGTCAAG 1931
DB GGAAGTATTGTACTCCAGCACTTTGGGAGGCCAAGGAGTGGATCACTCAGGTCAAG 86313
QY 1932 AGTTCAAAACCAAGCTGGCCACATGGTGAACCCCGCTCTCTACTAAATAACAAAAAT 1991
DB AGTTCTGAGACCAAGCTGGCCAAATGGTGAACCTCGTCTCTACTAAATAATAC-AAAT 86254

QY 1992 TAGCTAAGTGTGTGGCGCATGCTGTAATCCAGCTACTTTGGAGGGTGAAGGAGGAGA 2051
DB TAGCCAGGATGTGTGGCGGTGCTGTAAATCCAGCTACTCTTGAGGCTGAGGAGGAGA 86194
QY 2052 ATTTCTTTGAACCCCGGAGGAGAGGTTGCAGTGAAGCGAGATCACACACTGCACTCCAG 2111
DB ATTCCCTGAACCCCGGAGGAGGAGGTTGCAGTGAAGCGATATCGCAGCTGCACTCCAG 86134
QY 2112 CTTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGAAATTAAGCAAT 2171
DB CCTGGCGCAGAGAGTGAAGTCCGTCTCAAAAAACAAAAACAAAAAGATGTTTGTACTTCAA 86074
QY 2172 TAGCATTTCGAGA 2184
DB TGGCACTTAAAGA 86061
RESULT 11
US-09-735-934A-3
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: LI, Jiaxin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00851
; CURRENT APPLICATION NUMBER: US/09/735.934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 7.9%; Score 307.6; DB 4; Length 43950;
Best Local Similarity 67.6%; Pred. No. 4e-66;
Matches 491; Conservative 0; Mismatches 229; Indels 6; Gaps 4;
QY 1442 TTTTTTTTTTTTTTTTGTGAGAAAGAGTCTTGTCTCTTTTCACTCCCTGGCTGGAGGCGAGTGT 1501
DB TTTTTTTTGTGTGTGTGAGACAGAGTCTTGTCTCCGTCACTCCAGAGTGGAGTGCAGTGGCGT 10021
QY 1502 GATCTCAGCTCAGTGCACACTTTGCGCTCCCGGGTTCAAGCAATCTCTCTGCTTCAAGCTC 1561
DB GGTCTTGGCTCAGTGCACACTCTACCTCTCTGGGTTCAAGCAATCTCTCTGCTTCAAGCTC 10081
QY 1562 CCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTGTAGT 1621
DB CCAATAGCTGGGATTTACAGTGGCGCACCACTCTCTGGCTGATTTTGTATTTTGTAGT 10141
QY 1622 AGAGATGGGATTTTCACTTTTGTGGCAAGCTGCTCTCAAACTTTTGTGTCATAATTTGT 1681
DB --AGACGGGGTTTCCACCATGTTGGCCAGCTGGTTTCAAGTAATCCACCTCTCAGCT 10199
QY 1682 TGTAACTATTGTTCTTTTGTGTAGTAGGCGGCCACACCAACCAAAATAAATTTTGA 1741
DB CCCCAGTGTGGGATTTACAGCATGAGCCACTGGCTGGGCCCCCATCGGCTTCCATTTT 10259
QY 1742 ATCCAAATCAGTGTGTTGTTTGTACCACTGCTTGTAGAACACAGTGTGACAGGAGCC 1801
DB TGTATGTTGCTCCCTGCAATTTAGACCATATTTCTTGGATGTTTCCATTTGGGTATTAGTCT 10319
QY 1802 TCAGGATGAGGATGATCTCTGCTCGAAGAGAAATAGAAATGAAATATTTCTCGGGCCA 1861
DB GAGCAGCATCTTAGCTCCGTGGGTGCCACGCTTGTACAGAAATCCCTGATTTCTGGCCA 10379
QY 1862 GCGTGGTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCCAAGGATGAGGATCACC 1921
DB GGCACGGTGGCTCACACCTGTAATCCAGCACTTTGGAGGCCAAGGCGGGGATCA-- 10437


```

Db 50106 -----GCATTGGTTACCTTTTATGGTTGGCACT--GGATGGATGGATGGCCATTTA 50058
Qy 1800 CCTCAGGAGTAGAGTGTATCTCTGCTCGAAAGAAATAGAAATGAAATATCTTCGGGC 1859
Db 50057 TGGTAGTGATGCAACATTTCTCTCAAAATAAAATTAAGTCAGTTTAAAGGTAAAGCCAGGC 49998
Qy 1860 CAGCGTGGTGGCTATGCTGATCCAGCATTTCAGCAGCATTTCGGAGGCCAAGCATGTGGATCA 1919
Db 49997 CGGCGACGGTGGCTCAGCTTGTATCCAGCAGTTTGGGAGGGTGGAGCGAGTGGGAACA 49938
Qy 1920 CTTGAGTCAAGGAGTTCAAACACAGCTGGCCACATGGTGAAACCCCGTCTCTACTAAA 1979
Db 49937 CTTGAGTCAAGGAGTTCAGACCGAGCTGACCAACATGGAGAAACCCCGTCTCTACTAAA 49878
Qy 1980 AATACAAAAAATTAGCTAAGTGTGGTGGCGCATCCCTGTAAATCCAGCTACTTGGGAGGG 2039
Db 49877 AATAC--AAAAATTAGCCAGGTGTGGTGGCGCATGAACGTAATCCAGCTACTCCTGGGAAGC 49820
Qy 2040 TGAGCGAGGAGAAATTTCTTGAACCCGGAGGCAGAGGTTGCAGTGAAGCGAGATCAACC 2099
Db 49819 TGAGCGAGGAGAAATCGGTTTGAACCCAGGAGGCAGAGGTTGGGAGGAGCAAGATCGTGCC 49760
Qy 2100 ACTCACTCCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2156
Db 49759 ATTGCATCTCAGCTGGGCAACAAAGCGAAACTCTCTCAAAATAAATAAATAA 49702

RESULT 14
US-09-820-924-3/c
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
US-09-820-924-3

Query Match 7.8%; Score 304; DB 4; Length 39982;
Best Local Similarity 56.3%; Pred. No. 3e-65;
Matches 829; Conservative 0; Mismatches 595; Indels 48; Gaps 12;

Qy 735 TATTTTTTTTCTGAAACAGGGTATCACTCTGTTGCCAGGCTGGAGTACAGTGGCGTAA 794
Db 32734 TATTAATTTTTTTCAGATGGAGTCTCACTTTGTGCGCCATGCTGGAGTGCAGTGACGTGA 32675
Qy 795 TAATAGTCACTGAGCGTCC-CTCTGGGCTCAAGCAATCCGCTGGCTCAGCATCCT 853
Db 32674 TCTTGGCTCACTAACAACCTCCACCTCCCGGGTTCAAGCTATTTCTGTGCTCAGCCCTCCC 32615
Qy 854 GAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCAGCTAAGTTTAAATAATGATTTT 913
Db 32614 GAGTAGCTGGGATACAGGCAACCCACACACCACTGCTGGCTAAATTTTT-----GTACTTT 32561
Qy 914 TGGTATAGAGGAGTCTTGTCTATGTTGTCTCAGGCTGTATTTTTTATGTTGAGACAAGTTC 973
Db 32560 TAGTAGAGATGGGTTTCAACCATGTTGGCCGCGCTGGCTTGAATCTGTGCTCAAGTGA 32501
Qy 974 TCATATGTTGCCATGATCCCGCCACCTCCATCTTCCCAAGATGCTCATC--TTATCTGTT 1031
Db 32500 TCCGCTCTTGGCTCCCAAGTGTGTGGGATACAGGTGTGTGCCACACGCCCGCAGCT 32441
Qy 1032 CATTAGTCAGTTGACAGACATTTAGTTGTTTCCACTTTTGGACCATATGAATAACT 1091

```

```

Db 32440 AATTTTGTATTTTTTAGTGAGAAAGGGTTTCAACATGTTGGCCAGGCTGCTCTCGAACT 32381
Qy 1092 CCAGTGAATATTCATGTANACATTTGTGTGGGCATATGTTTTCAATT----- 1138
Db 32380 CTTGACCTCGTGTATTCGGCGCCTCAGCCTCCCAAAGTGTGGGATTAACAAGCCAGGCC 32321
Qy 1139 -----CTGTTTGGTATTATCTAGGATGGAATGTCTGGATCCCGGTAATATATTTTGACA 1193
Db 32320 TGATACAGTGGTTTAAAAATGAGGAAGATCTTCAAAACCAATGTGGAATCATTTTCTTA 32261
Qy 1194 GGAGAGTTTCAGGGGAGAAAAAATCTTGGGAAAAATGAAGCATGTTTAG-----AAA 1243
Db 32260 ATAAATTTTAACGAAGAAATGCAAAAGGAAAGCAGAGTAGTCTAGGGTACTACTTCTT 32201
Qy 1244 TCAGCAAGAGTGCAGGGTTTTTCGGAGTTTTTATTTATATTTCTGTTGACAAATGTGCGAG 1303
Db 32200 GCTGTGAGAAAGGGAAGATAGCATACTTGTCTGCTGTTTGTGCTAAACGAAAC 32141
Qy 1304 TTTGATGAAGATACAAGTTATCTAAGTGAGAGTGAAGATTAAGGCTGGAATAGGGCGT 1363
Db 32140 CAGGAGGAGGAACAG--AAGTGAACGAGACTGCTTACCTACAGGGGGTAGGGGAGA 32083
Qy 1364 TCAGAGTAAATCATGAAGCACTTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAA 1433
Db 32082 AAACAGGCAAAAGGAGAGTTACTGTATTGTATAGTTCTGACTTTGGAGCTACATTAA 32023
Qy 1424 ATAAATAAAATCACAATTTTTTTTTTTTTTTTTTTTGGAGAAAGTCTTCTCTTTCACCTG 1483
Db 32022 TGTCTTTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 31964
Qy 1484 GCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTTCGCTCCCTCC--GGTTTCAAGCA 1542
Db 31963 GCTGGAGTGCAGTGTGTGATCAAGCTCACTGAGCTCGACTCCAGGGCTCAGGTG 31904
Qy 1543 ATTCTCTGCTTCCAGCTCCCAAGTAGCTGGGACTCAGGCACTTCCACCATGCCCAGC 1602
Db 31903 ATCTCCCACTCAGCTCCAGCTCAGGTAGCTGGGACTCAGGTGCACTACACACCCAGC 31844
Qy 1603 TGATTTTGTATTTTATGTAGAGATGGGATTTACATTTGTTGGCCAGCTGCTCTCAAC 1662
Db 31843 TAAATTTTGTATTTCTTGTAGACA--GAGGTTTTCGCCATGTTGCCAGGCTGGTCTCAAC 31785
Qy 1663 T-TTTTGTCTGTCATAATTGTGTAACTTATTTGTTCTTTTGTGAGGTAGGCGCCCCAGAC 1721
Db 31784 TCCTGGGCTCAAGCAATCCATCTGCTTGGCTCCCAAAGTGTAGGATTAACAGGCATGA 31725
Qy 1722 CAAAAAATAAATCTTTAGAAATCCAAATCAGTGTGTTGTTGACCACTGTCTCACTTGAGA 1781
Db 31724 GACACTGTGCCAGCCACGTTGTTTCAAAATTCAAAAACAACAGCAACAATAACAAACG 31665
Qy 1782 ACCACAGTGTGACAGGGCTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAA 1841
Db 31664 CTCAAAA CCAAAACACAAAATCAGAAATGAGAGGGAATCCCTAAAGTGGAAATAAAA 31605
Qy 1842 TGAATAATTTCTCGGCGCCAGGGTGTGTGCTCATGCTGTATTCAGCAGCATTTGGGAG 1901
Db 31604 TGATCCAA---CTAGGCTGGCGCGGTGCTCACACTGTATTCAGCAGCATTTGGGAG 31548
Qy 1902 GCCAAGGCATGTGATCACTGAGGTGAGGTTCAAAAACAGCTGCGCAACATGTGTA 1961
Db 31547 GCCGAGCGGCGGATCAC--GAGGTGAGAGATCGAGACCATCTCTGGCTAAACCATGA 31490
Qy 1962 AACCCGCTCTACTAATAAATCAAAAAATAGTAACTGTGGTGGCGCATGCTGTAAT 2021
Db 31489 AACCCCACTCTACTAATAAATCAAAAAATAGCCGGCGGTGGTAGCGGCACCTGTAGT 31430
Qy 2022 CCCAGCTACTTGGAGGCTGAGGCGAGGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGCA 2081
Db 31429 CCCAGCTACTTGGAGGCTGAGGCGAGGAGATGCGGTGAACCCCGGAGGAGAGGTTGCA 31370
Qy 2082 GTGAAGCGAGATCACCACTGCACTCAGCTCGGGGAGAGAGGAGACTTCTCTCAA 2141
Db 31369 GTGAGCTGAGATAGTGCCACTTGCACTCCGGCTGGGCGAAAGAGCAAGACTCCGCTCTCAA 31310

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:06:53 ; Search time 2243.23 Seconds
(without alignments)
7909.918 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910

Perfect score: 3910

Sequence: 1 aggaattcatcatttaaat.....cggaacaggtcacccgtgc 3910

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
 - 17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
 - 19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3910	100.0	4150	14	US-10-016-725-15
2	3881.8	99.3	37698	13	Sequence 15, Appl
3	630.6	16.1	631	13	Sequence 10, Appl
4	630.6	16.1	631	16	Sequence 24494, A
5	460	11.8	59725	13	Sequence 24494, A
6	450.4	11.5	617	10	Sequence 814, App
7	407.4	10.4	9631	15	Sequence 1, Appl
8	360.4	9.2	23618	13	Sequence 1569, Ap
9	353.6	9.0	36296	13	Sequence 4, Appl
10	353	9.0	60815	13	Sequence 1584, Ap
11	349.8	8.9	12394	16	Sequence 52, Appl
12	349	8.9	13216	10	Sequence 10, Appl
13	347.4	8.9	29163	10	Sequence 7810, Ap
14	342.8	8.8	26928	9	Sequence 7809, Ap
					Sequence 2278, Ap

15	342.8	8.8	26928	13	US-10-235-192A-34
16	342.8	8.8	26928	15	US-10-020-141-7
17	342.8	8.8	26928	15	US-10-017-631-1
18	338.6	8.7	144035	13	US-10-087-192-322
19	337.2	8.6	3273	16	US-10-012-697-1406
20	337	8.6	109906	13	US-10-235-192A-31
21	334.8	8.6	108182	13	US-10-087-192-1618
22	333.8	8.5	175077	13	US-10-087-192-1168
23	333.6	8.5	35425	15	US-10-017-161-2429
24	333.6	8.5	35425	16	US-10-292-798-2069
25	332.2	8.5	3158	13	US-10-027-632-115030
26	332.2	8.5	3158	16	US-10-027-632-115030
27	330.6	8.5	3158	13	US-10-027-632-115030
28	330.6	8.5	3158	16	US-10-027-632-115030
29	330.2	8.4	198522	13	US-10-087-192-244
30	329.6	8.4	9372	9	US-09-764-877-3049
31	329.6	8.4	9372	9	US-09-764-877-3244
32	329.6	8.4	9372	16	US-10-242-515-3049
33	329.6	8.4	9372	16	US-10-242-515-3244
34	328.8	8.4	2130	16	US-10-108-260A-2246
35	328.8	8.4	75252	13	US-10-087-192-904
36	328.6	8.4	35425	16	US-10-292-798-2069
37	328.6	8.4	35425	16	US-09-997-722-154
38	328.4	8.4	96594	13	US-10-087-192-844
39	328	8.4	91760	13	US-10-087-192-844
40	327.6	8.4	12919	15	US-10-017-161-787
41	327.6	8.4	17397	9	US-09-764-869-1945
42	327.6	8.4	17397	15	US-10-091-504-1945
43	327.6	8.4	17397	16	US-10-227-577-1945
44	327.6	8.4	19334	9	US-09-764-869-1943
45	327.6	8.4	19334	15	US-10-091-504-1943

ALIGNMENTS

RESULT 1

US-10-016-725-15

; Sequence 15, Application US/10016725

; Publication No. US20020151018A1

; GENERAL INFORMATION:

; APPLICANT: Prouty, Stephen

; APPLICANT: Zhang, Lin

; APPLICANT: Stearn, Kurt

; TITLE OF INVENTION: Stearyl-CoA Desaturase Promoter

; FILE REFERENCE: J622065

; CURRENT APPLICATION NUMBER: US/10/016,725

; CURRENT FILING DATE: 2002-06-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 15

; LENGTH: 4150

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-016-725-15

Query Match 100.0%; Score 3910; DB 14; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGGAATTCATCCATTTAAATCATACAAATTAATGSCITTTTAGTATATTCACAGGTTGTC	60
DB	1	AGGAATTCATCCATTTAAATCATACAAATTAATGSCITTTTAGTATATTCACAGGTTGTC	60
QY	61	ATCCATCAATCCATTTTAGAACGTTTTTACTCCAAAATAAACCCCTGATTCCTT	120
DB	61	ATCCATCAATCCATTTTAGAACGTTTTTACTCCAAAATAAACCCCTGATTCCTT	120
QY	121	AGCCATACCCCCCAACATCTCCATCTCTTCCAGCCCTGGGCAACCCCAATCTAC	180
DB	121	AGCCATACCCCCCAACATCTCCATCTCTTCCAGCCCTGGGCAACCCCAATCTAC	180
QY	181	TTTCTGCTCTATTAATTTGCCAATTCCTGGACATTTATTAATGGAAGCAACAT	240

181	Db	 TTTCTGCTCTATAAAATTTGCCAAATTTCTGGACATTTTCATATAAATGGAAGCAAAACAACAT	240
241	Qy	GTGAGACTTTGTGACATCGGCTGCTTTCACATAGCATTTCTATTTTTTAAGGCTCATATATGTTA	300
241	Db	GTGAGACTTTGTGACATCGGCTGCTTTCACATAGCATTTCTATTTTTTAAGGCTCATATATGTTA	300
301	Qy	CAGTACTTAGCAGCATCTTCATTTCTTTTTTATTTCTCAAATGGTATTTCCATGTGTGGGTAT	360
301	Db	CAGTACTTAGCAGCATCTTCATTTCTTTTTTATTTCTCAAATGGTATTTCCATGTGTGGGTAT	360
361	Qy	CCCATATCATATTTATGAGACAGGTTCTCACTCTGTCAACCAGGCTGAGTGCAGTGGC	420
361	Db	CCCATATCATATTTATGAGACAGGTTCTCACTCTGTCAACCAGGCTGAGTGCAGTGGC	420
421	Qy	ACAAATCATAGCTCACATGTAACTCCTCGGGCTCAAGTGCATCTCACTACTCCTCAGGC	480
421	Db	ACAAATCATAGCTCACATGTAACTCCTCGGGCTCAAGTGCATCTCACTACTCCTCAGGC	480
481	Qy	TCCAGAGTAGCTAGGACTACAGGCACACACAGGCCATACCTGGCTAAATTTTTTTTTTAAT	540
481	Db	TCCAGAGTAGCTAGGACTACAGGCACACACAGGCCATACCTGGCTAAATTTTTTTTTTAAT	540
541	Qy	TTTCAATTTATGATATTCATTTCTTTCTTTTTTTTGTGTGTGTGTGTGTGTGATAGGTCCTC	600
541	Db	TTTCAATTTATGATATTCATTTCTTTCTTTTTTTTGTGTGTGTGTGTGTGTGATAGGTCCTC	600
601	Qy	ACTTTGTATCCAGGCTGGAGGCGAGTGGCATGCTGACAGCTCAGCAGGCTTCACATTCCT	660
601	Db	ACTTTGTATCCAGGCTGGAGGCGAGTGGCATGCTGACAGCTCAGCAGGCTTCACATTCCT	660
661	Qy	GGGCTCAAGTGATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAAACGCTGTCA	720
661	Db	GGGCTCAAGTGATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAAACGCTGTCA	720
721	Qy	CCATGCTGGCTGATATTTTTTTTTTTCTTTGAAACAGGGTATCACTCTGTGTCGCCAGGCTGA	780
721	Db	CCATGCTGGCTGATATTTTTTTTTTTCTTTGAAACAGGGTATCACTCTGTGTCGCCAGGCTGA	780
781	Qy	GTACAGTGGGGTAAATATAGCTCACTGCAGCCTCCCTCCTGGGCTCAAGCAATCCGCTG	840
781	Db	GTACAGTGGGGTAAATATAGCTCACTGCAGCCTCCCTCCTGGGCTCAAGCAATCCGCTG	840
841	Qy	GCCTCAGCATCCTCAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCACAGCTAAATTTT	900
841	Db	GCCTCAGCATCCTCAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCACAGCTAAATTTT	900
901	Qy	AAAAAATGATTTTTTGGTATAGAGAGGTCCTTGCTATGTTGCTCAGGCTGTATTTTTATTG	960
901	Db	AAAAAATGATTTTTTGGTATAGAGAGGTCCTTGCTATGTTGCTCAGGCTGTATTTTTATTG	960
961	Qy	TTGAGACAGGCTCTCACTATGTTGCCATGATCCCCCACCCTCCACTTCCCAAGTGTCTCA	1020
961	Db	TTGAGACAGGCTCTCACTATGTTGCCATGATCCCCCACCCTCCACTTCCCAAGTGTCTCA	1020
1021	Qy	TCTTATCTGTTCAATTTAGTTCAGTTCACACAGATTTTAGTGTGTTCCACTTTTTGACCAATTA	1080
1021	Db	TCTTATCTGTTCAATTTAGTTCAGTTCACACAGATTTTAGTGTGTTCCACTTTTTGACCAATTA	1080
1081	Qy	TGAATAATPACTCCAGTGAATATTCATGTATACATTTGTGTGGGCAWAGTTTTTCATTTCT	1140
1081	Db	TGAATAATPACTCCAGTGAATATTCATGTATACATTTGTGTGGGCAWAGTTTTTCATTTCT	1140
1141	Qy	GTGTGGTTTTATATCTAGAGTGGAATGCTCGAATCCGGGTAAATATTTTTCACAGGCGAG	1200
1141	Db	GTGTGGTTTTATATCTAGAGTGGAATGCTCGAATCCGGGTAAATATTTTTCACAGGCGAG	1200
1201	Qy	TTTCAGGGGAAGAAAAAATTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGG	1260
1201	Db	TTTCAGGGGAAGAAAAAATTTGGGAAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGG	1260
1261	Qy	GTTTTTTCGGAGTTTTATTTTTATATTTCTGTTCACAAATGTGCAGTTTGTATGAAGATACAAG	1320

1361	Db	GT	TTTTTCGGAGTTTTTA	TTTTTA	TCTGTGTGA	CAAA	TGTCGAGTTT	GATCGAAGTACAAG	1321	
1321	QY	TT	TACTAA	GTGAGAGTGAGAA	TAAAGCGT	TGGAATA	TAGGGCGTT	TCAGAGTAAATCATGA	1381	
1321	Db	TT	TACTAA	GTGAGAGTGAGAA	TAAAGCGT	TGGAATA	TAGGGCGTT	TCAGAGTAAATCATGA	1380	
1381	QY	AG	CACTTTGA	ATACCAAAATTA	AGGAGCTT	GGCTGT	THAAA	CAAAAATATATAAAATACAA	1440	
1381	Db	AG	CACTTTGA	ATACCAAAATTA	AGGAGCTT	GGCTGT	THAAA	CAAAAATATATAAAATACAA	1440	
1441	QY	TTTTTTTTTTTTTTTT	TTTTTCAG	AAAGAGCTT	CGCTCT	TTTTC	CA	CCCTGGCTGGAGGCGAGTGGTG	1500	
1441	Db	TTTTTTTTTTTTTTTT	TTTTTCAG	AAAGAGCTT	CGCTCT	TTTTC	CA	CCCTGGCTGGAGGCGAGTGGTG	1500	
1501	QY	TG	ATCTCAG	CTCACTGCA	ACTTTT	CGCCTCC	GGGTTCA	AGCAAAATPCCTCTGCTTCAGCCT	1560	
1501	Db	TG	ATCTCAG	CTCACTGCA	ACTTTT	CGCCTCC	GGGTTCA	AGCAAAATPCCTCTGCTTCAGCCT	1560	
1561	QY	CC	AAAGTAG	CTGGGACTAC	AGGCAT	TTCC	CA	CANVCCAGCTGATTTTGTATTTTAT	1620	
1561	Db	CC	AAAGTAG	CTGGGACTAC	AGGCAT	TTCC	CA	CANVCCAGCTGATTTTGTATTTTAT	1620	
1621	QY	TAG	AGATGGGA	ATTTCA	CTTTGTT	GGCCAA	GTGCT	CTCAAACTTTTTTGCTCTCATATTTG	1680	
1621	Db	TAG	AGATGGGA	ATTTCA	CTTTGTT	GGCCAA	GTGCT	CTCAAACTTTTTTGCTCTCATATTTG	1680	
1681	QY	TT	TAAC	TATTTG	TCTTTT	GCTGAG	TAGGG	CCCCCAGACCAAAAAATATAATCTTTAG	1740	
1681	Db	TT	TAAC	TATTTG	TCTTTT	GCTGAG	TAGGG	CCCCCAGACCAAAAAATATAATCTTTAG	1740	
1741	QY	AAT	CCAAAT	CAGTGTGT	TTGATTT	GAC	ACTGT	CTACTTGAGAACACAGTGTGAC	1800	
1741	Db	AAT	CCAAAT	CAGTGTGT	TTGATTT	GAC	ACTGT	CTACTTGAGAACACAGTGTGAC	1800	
1801	QY	CT	CAGGAG	TAGAGT	TGATCT	CTGCT	CGAAAGAG	AAAAATAGAAATATAATTTCTCCGGGCC	1860	
1801	Db	CT	CAGGAG	TAGAGT	TGATCT	CTGCT	CGAAAGAG	AAAAATAGAAATATAATTTCTCCGGGCC	1860	
1861	QY	AG	GGTGTG	TGCTCAT	GCTGTAA	TCC	CAGCA	CTTTGGGAGGCCAAGGCATGTGGATCAC	1920	
1861	Db	AG	GGTGTG	TGCTCAT	GCTGTAA	TCC	CAGCA	CTTTGGGAGGCCAAGGCATGTGGATCAC	1920	
1921	QY	CT	GAGG	CTCAG	AGTTCA	AAAC	CCAGCT	TGGCCAA	CATGGTGA	1980
1921	Db	CT	GAGG	CTCAG	AGTTCA	AAAC	CCAGCT	TGGCCAA	CATGGTGA	1980
1981	QY	AT	ACAAAAAT	TAGCTA	AGTGTGT	GCGCAT	GCTGTAA	TCC	CAGCTACTTTGGGAGGGT	2040
1981	Db	AT	ACAAAAAT	TAGCTA	AGTGTGT	GCGCAT	GCTGTAA	TCC	CAGCTACTTTGGGAGGGT	2040
2041	QY	GAG	G	CAGAGAA	TTTCTTTGA	CCCCGG	AGCAGAGT	TG	CAGTGAACGAGATCACACCA	2100
2041	Db	GAG	G	CAGAGAA	TTTCTTTGA	CCCCGG	AGCAGAGT	TG	CAGTGAACGAGATCACACCA	2100
2101	QY	CT	GACTCC	AGCTTGGGG	GAGAGCGAG	ACTT	CCTCT	CA	AAAAAACA	2160
2101	Db	CT	GACTCC	AGCTTGGGG	GAGAGCGAG	ACTT	CCTCT	CA	AAAAAACA	2160
2161	QY	AT	TAGCAAA	TTAGAC	ATTG	CAGAGAGAA	CCTG	AAAGGGGGT	TCAGACCA	2220
2161	Db	AT	TAGCAAA	TTAGAC	ATTG	CAGAGAGAA	CCTG	AAAGGGGGT	TCAGACCA	2220
2221	QY	TG	TGCCA	TATGCC	AA	GTACTT	CTGAGG	CATG	ACTGATGAGCTGTCCACATCTGAAATCA	2280
2221	Db	TG	TGCCA	TATGCC	AA	GTACTT	CTGAGG	CATG	ACTGATGAGCTGTCCACATCTGAAATCA	2280
2281	QY	TC	CAGTCT	TTTTCAGAA	CTTTTCA	CACCGG	ACAGGG	ACCGAG	CTGGNATCGACTCTCCT	2340
2281	Db	TC	CAGTCT	TTTTCAGAA	CTTTTCA	CACCGG	ACAGGG	ACCGAG	CTGGNATCGACTCTCCT	2340
2341	QY	GT	CTACTGG	CCAGAGTT	TG	CCCT	TGAG	CAC	TAGTGGCCAA	2400
2341	Db	GT	CTACTGG	CCAGAGTT	TG	CCCT	TGAG	CAC	TAGTGGCCAA	2400

183 TCCTGCTCTATAAATTGGCAATTTCTGGACATTTCTATATAAATGGAAGCAAAACAATGT 242
 Db
 5460 TCCTGCTCTATAAATTGGCAATTTCTGGACATTTCTATATAAATGGAAGCAAAACAATGT 6519
 QY
 243 GAGACTTTGTGACTGGCTGCTTTTCACTTAGCATTTCTATATTTTAAAGGCTCATATATGTACA 302
 Db
 6520 GAGACTTTGTGACTGGCTGCTTTTCACTTAGCATTTCTATATTTTAAAGGCTCATATATGTACA 6579
 QY
 303 GTACTTAGAGTACTTCAATTTCTTTTATTTCTCAATGGTATTTCCACTGTGTGGGTATCC 362
 Db
 6580 GTACTTAGAGTACTTCAATTTCTTTTATTTCTCAATGGTATTTCCACTGTGTGGGTATCC 6639
 QY
 363 CATATCATATTTATTAGAGACAGGTTCCTCACTGTCTCACCCAGGCTGGAGTGCAGTGGCAC 422
 Db
 6640 CATATCATATTTATTAGAGACAGGTTCCTCACTGTCTCACCCAGGCTGGAGTGCAGTGGCAC 6699
 QY
 423 AATCATAGCTCACTGTAACTTCAAACTCTCTGGGCTCAAGTGAATCTTACTACCTCAGGCTC 482
 Db
 6700 AATCATAGCTCACTGTAACTTCAAACTCTCTGGGCTCAAGTGAATCTTACTACCTCAGGCTC 6759
 QY
 483 CAGAGTAGCTAGGACTACAGGCAACACAGCCATACCTGGCTAATTTTATTTTATTTT 542
 Db
 6760 CAGAGTAGCTAGGACTACAGGCAACACAGCCATACCTGGCTAATTTTATTTTATTTT 6819
 QY
 543 TCATTTTATGTAATTTCTTTTCTTTTGTGTTGTTTGTAGATAGGCTCTCAC 602
 Db
 6820 TCATTTTATGTAATTTCTTTTCTTTTGTGTTGTTTGTAGATAGGCTCTCAC 6879
 QY
 603 TTTGTTTACCGAGCTGAGGAGCTGGTGAAGTGTGACAGCTGAGCAGCTTTGACTTCTCTGG 662
 Db
 6880 TTTGTTTACCGAGCTGAGGAGCTGGTGAAGTGTGACAGCTGAGCAGCTTTGACTTCTCTGG 6939
 QY
 663 GCTCAAGTGTATCTCTGCTCCAGCTCCCAAGTGTGGGACTACAAACACAGTGTCAAC 722
 Db
 6940 GCTCAAGTGTATCTCTGCTCCAGCTCCCAAGTGTGGGACTACAAACACAGTGTCAAC 6999
 QY
 723 ATGCTCTGGCTGATATTTTCTTTGAAACAGGGTATCACTCTGTGTCGCCAGGCTGGAGT 782
 Db
 7000 ATGCTCTGGCTGATATTTTCTTTGAAACAGGGTATCACTCTGTGTCGCCAGGCTGGAGT 7059
 QY
 783 ACAGTGGGTAATTAATAGTCACTGTGACCTCCCTCTCTGGGCTCAAGCAATCCGCTGGC 842
 Db
 7060 ACAGTGGGTAATTAATAGTCACTGTGACCTCCCTCTCTGGGCTCAAGCAATCCGCTGGC 7119
 QY
 843 CTACAGTATCTGAGTGTGGGACTACAGGCTGTGGCCACAGGCCCCAGTAACTTATTA 902
 Db
 7120 CTACAGTATCTGAGTGTGGGACTACAGGCTGTGGCCACAGGCCCCAGTAACTTATTA 7179
 QY
 903 AAAATGATTTTGGTATAGAGAGGTCTGTGCTATGTTGCTCAGGCTGTATTTTATTTGTT 962
 Db
 7180 AAAATGATTTTGGTATAGAGAGGTCTGTGCTATGTTGCTCAGGCTGTATTTTATTTGTT 7239
 QY
 963 GAGACAGGCTCTACTATGTTGCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCATC 1022
 Db
 7240 GAGACAGGCTCTACTATGTTGCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCATC 7299
 QY
 1023 TTATCTGTTTCAATAGTGTAGTGTAGACATTTAGGTTGTTTCCACTTTTGTGACCAATATG 1082
 Db
 7300 TTATCTGTTTCAATAGTGTAGTGTAGACATTTAGGTTGTTTCCACTTTTGTGACCAATATG 7359
 QY
 1083 AATAATATCTCAGTGAATATTAATGATATACATTTGTTGTTGGGATATGTTTTCATTTCTGT 1142
 Db
 7360 AATAATATCTCAGTGAATATTAATGATATACATTTGTTGTTGGGATATGTTTTCATTTCTGT 7419
 QY
 1143 TGGGTTTATATCTAGGAGTGAATGCTGTGATCCCGGTTATATTTTGCACAGGAGAGTT 1202
 Db
 7420 TGGGTTTATATCTAGGAGTGAATGCTGTGATCCCGGTTATATTTTGCACAGGAGAGTT 7479
 QY
 1203 CAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGT 1262
 Db
 7480 CAGGGGAAGAAAACCTTGGGAAAATGAAGCATGTTTGAAGAAATCAGCAAGAGTGCAGGGGT 7539
 QY
 1263 TTTTCGGAGTTTATTTATTTATTTCTGTTTGAACAAATGTGCAAGTTTGTGATGAAGATCAAGTT 1322

7540 TTTTCGGAGTTTATTTTATTTCTGTTGACAAATGTGCAAGTTTGTGATGAAGATCAAGTT 7599
 QY
 1323 ATACTTAAGTGAAGCTGAGAAATTAAGGCTGGATAGGCGTTTCAGAGTAAATCATGAAG 1382
 Db
 7600 ATACTTAAGTGAAGCTGAGAAATTAAGGCTGGATAGGCGTTTCAGAGTAAATCATGAAG 7659
 QY
 1383 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAAAAATCAAAAT 1442
 Db
 7660 CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAAATAAATAAAAAATCAAAAT 7719
 QY
 1443 TTTTCTTTTGTGAGAAAGAGTCTTGTCTTTTACCCTGTGAGGCGAGTGTGTG 1502
 Db
 7720 TTTTCTTTTGTGAGAAAGAGTCTTGTCTTTTACCCTGTGAGGCGAGTGTGTG 7779
 QY
 1503 ATCTCAGCTCACTGCAACTTTTCCCTCCCGGTTCAAGCAATTTCTCTGCTTCAGGCTCC 1562
 Db
 7780 ATCTCAGCTCACTGCAACTTTTCCCTCCCGGTTCAAGCAATTTCTCTGCTTCAGGCTCC 7839
 QY
 1563 CAAGTAGCTGGGACTACAGGCACTTCCCACTGCCCAGCTGATTTTGTATTTTAGTA 1622
 Db
 7840 CAAGTAGCTGGGACTACAGGCACTTCCCACTGCCCAGCTGATTTTGTATTTTAGTA 7899
 QY
 1623 GAGATGGATTTTCACTTTTGTGCGCAAGCTGTCTCAACTTTTGTCTGTCAATATGTT 1682
 Db
 7900 GAGATGGATTTTCACTTTTGTGCGCAAGCTGTCTCAACTTTTGTCTGTCAATATGTT 7959
 QY
 1683 GTAACATTTGTTTCTTTTGTGAGGCTAGGCGCCCAAGCAAAATAAAATCTTAGAA 1742
 Db
 7960 GTAACATTTGTTTCTTTTGTGAGGCTAGGCGCCCAAGCAAAATAAAATCTTAGAA 8019
 QY
 1743 TCCAAATCAGTGTGTTTGTGACCACTGTCTGAGAAACCAAGTGTGACAGGCGCT 1802
 Db
 8020 TCCAAATCAGTGTGTTTGTGACCACTGTCTGAGAAACCAAGTGTGACAGGCGCT 8079
 QY
 1803 CAGAGTAGAGTGTCTCTGCTGAAAGAGAAATAGAAATGAAATTTCTCGGGCCAG 1862
 Db
 8080 CAGAGTAGAGTGTCTCTGCTGAAAGAGAAATAGAAATGAAATTTCTCGGGCCAG 8139
 QY
 1863 GCGTGTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCCCAAGGCAATGTGATCACT 1922
 Db
 8140 GCGTGTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCCCAAGGCAATGTGATCACT 8199
 QY
 1923 GAGTCAAGAGTTTCAAAACCAAGCTGCTCAAAATGTGAAACCCGCTCTCTACTAAAAAT 1982
 Db
 8200 GAGTCAAGAGTTTCAAAACCAAGCTGCTCAAAATGTGAAACCCGCTCTCTACTAAAAAT 8259
 QY
 1983 AAAAAAATTAGCTTAAGTGTGGGCGCATGCTGTAAATCCAGCTACTTTGGGAGGGTGA 2042
 Db
 8260 AAAAAAATTAGCTTAAGTGTGGGCGCATGCTGTAAATCCAGCTACTTTGGGAGGGTGA 8319
 QY
 2043 GGCAGGAGAAATTTTGAACCCGGGAGGCGAGGTTGCAAGCGAGATCACACCACT 2102
 Db
 8320 GGCAGGAGAAATTTTGAACCCGGGAGGCGAGGTTGCAAGCGAGATCACACCACT 8379
 QY
 2103 GCCTCCAGCTCGGGGAGAGGCGAGCTTCTCTCAAAACCAAAACAAAAAGAAAT 2162
 Db
 8380 GCCTCCAGCTCGGGGAGAGGCGAGCTTCTCTCAAAACCAAAACAAAAAGAAAT 8439
 QY
 2163 TAGCAAAATTAGACATTTGACAGAGAACTTGAAGGGGCTCAGACCCAGTACAGATTTCTG 2222
 Db
 8440 TAGCAAAATTAGACATTTGACAGAGAACTTGAAGGGGCTCAGACCCAGTACAGATTTCTG 8499
 QY
 2223 TGGCACAATGCAAGTACTTCTGAGGCGATGCTGGAATGAGCTGTCCACATCTGAAATCATC 2282
 Db
 8500 TGGCACAATGCAAGTACTTCTGAGGCGATGCTGGAATGAGCTGTCCACATCTGAAATCATC 8559
 QY
 2283 CAGTCTTTTTCAGAACTTTTCAACCCGACAGGGAGCCAGGACTGGAATGAGTCTCTCTGG 2342
 Db
 8560 CAGTCTTTTTCAGAACTTTTCAACCCGACAGGGAGCCAGGACTGGAATGAGTCTCTCTGG 8619
 QY
 2343 TCCTGCGCAGAGAGTTTGGCTTGAACCTTGAACCTGAGACCTGCGCAACAAAGAGGCTGCTAG 2402

Db 8620 TCACTGGCCAGAGAGTGGCC-TGACCCCTGAGACCAAGTGGCAACAAAGAGAGCTGCTTAG 8678
Qy 2403 TCTACCTCCAGGAAATCCCAAGTGTCTTCTCTGGGAAGTGAATCAATTTGGCGCAGCA 2462
Db 8679 TCTACCTCCAGGAAATCCCAAGTGTCTTCTCTGGGAAGTGAATCAATTTGGCGCAGCA 8738
Qy 2463 CTCGTAATTTCTCTCTTCCAGGGAAGATCTTAGGGCAGTATTTGGGAAGACATG 2522
Db 8739 CTCGTAATTTCTCTCTTCCAGGGAAGATCTTAGGGCAGTATTTGGGAAGACATG 8798
Qy 2523 GGCATGAGAGACACCGGTTGAATGATAGCTGCTGCTTCTGAGCTCTCATGTTAAGG 2582
Db 8799 GGCATGAGAGACACCGGTTGAATGATAGCTGCTGCTTCTGAGCTCTCATGTTAAGG 8858
Qy 2583 CTCCTACAGACACGMAAAGATGGGGCAGAGGACAGATCAGTAGGTCAGAGCATCTC 2642
Db 8859 CTCCTACAGACACGMAAAGATGGGGCAGAGGACAGATCAGTAGGTCAGAGCATCTC 8918
Qy 2643 AGGACCGAGGGCAATATGCTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGTTGT 2702
Db 8919 AGGACCGAGGGCAATATGCTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGTTGT 8978
Qy 2703 TTCTGGGCTCACTGCGAGCTCGTCACTTACTGTTGCTGTGACCAATGGCAAGTTATT 2762
Db 8979 TTCTGGGCTCACTGCGAGCTCCATCACTTACTGTTGCTGTGACCAATGGGCAAGTTATT 9038
Qy 2763 CCATCTCTCCATATCTCTTCTCACTTTAAATGGAATAATGGGTACCCACCTCCCA 2822
Db 9039 CCATCTCTCCATATCTCTTCTCACTTTAAATGGAATAATGGGTACCCACCTCCCA 9098
Qy 2823 GGGTCACAGAGAGGCTTACAGAAACGATCTTGTGAATGGCTTGCAGTAAATTTCAA 2882
Db 9099 GGGTCACAGAGAGGCTTACAGAAACGATCTTGTGAATGGCTTGCAGTAAATTTCAA 9158
Qy 2883 TACCTGCCAGTATCTTATTTCCATCCAGCCCTTTCGCTGCTGCTGGTGAARACA 2942
Db 9159 TACCTGCCAGTATCTTATTTCCATCCAGCCCTTTCGCTGCTGCTGGTGAARACA 9218
Qy 2943 CATGTCACTGTTCTCTGACGGTTTCCAGAAAGATTTCCAAATTAACAACCTGCCAGTC 3002
Db 9219 CATGTCACTGTTCTCTGACGGTTTCCAGAAAGATTTCCAAATTAACAACCTGCCAGTC 9278
Qy 3003 TGAAGAATCTCCAAACATCCCGACCGATCTTGAGCGCGGGCTTGGGATGGGACTG 3062
Db 9279 TGAAGAATCTCCAAACATCCCGACCGATCTTGAGCGCGGGCTTGGGATGGGACTG 9338
Qy 3063 CCGGCCGGGTTCTGAACAGGATGGTGGCGCAGGCACACACACAGCCAGCTCTG 3122
Db 9339 CCGGCCGGGTTCTGAACAGGATGGTGGCGCAGGCACACACACAGCCAGCTCTG 9398
Qy 3123 TGTGGCGCGGAGTCCGGTGGCTCCGGGTGAGCAGCGCTGGTGGTGGCGGGCAG 3182
Db 9399 TGTGGCGCGGAGTCCGGTGGCTCCGGGTGAGCAGCGCTGGTGGTGGCGGGCAG 9458
Qy 3183 AGCCATTTTCGACGCGTACGAGCCCCCGGCTCGCCGGGAGGAGGGCGGGCTTC 3242
Db 9459 AGCCATTTTCGACGCGTACGAGCCCCCGGCTCGCCGGGAGGAGGGCGGGCTTC 9518
Qy 3243 CCGCGT-CCCCAAGTCCAGATCTTGGGTGCTGCAAGTCTCCTGTCACGCGCTGG 3301
Db 9519 CCGCGTCCCCAAGTCCAGATCTTGGGTGCTGCAAGTCTCCTGTCACGCGCTGG 9578
Qy 3302 GGGGAGCGGAACGAGGATTTAGTGTGGCGCCCGCCCGAGGGTTTCAACACTG 3361
Db 9579 GGGGAGCGGAACGAGGATTTAGTGTGGCGCCCGCCCGAGGGTTTCAACACTG 9638
Qy 3362 TTTCCTGAGAACTTCCCAAGTSCCAACCAACCGCTTCTCGGTGTCAGGCGCGGTC 3421
Db 9639 TTTCCTGAGAACTTCCCAAGTSCCAACCAACCGCTTCTCGGTGTCAGGCGCGGTC 9698
Qy 3422 CTGGGTAGGCTCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 3481
Db 9699 CTGGGTAGGCTCCGCGCCCGAGCCCAACCGGGTCCCGAGCCCTTCCAGAGAGAA 9758

Qy 3482 CTCCCGACGCGGATGCGCGGACAGGCCCAGCGCGGGTGGAAAGAGAGCTGAGAGGA 3541
Db 9759 CTCCCGACGCGGATGCGCGGACAGGCCCAGAGCGCGGGTGGAAAGAGAGCTGAGAGGA 9818
Qy 3542 GAAACAGAGGGGAGGAGAGAGCTGCGGCGAGAGGGAACAGCAGATTGCGCCGAG 3601
Db 9819 GAAACAGAGGGGAGGAGAGAGCTGCGGCGAGAGGGAACAGCAGATTGCGCCGAG 9878
Qy 3602 CCAATGCAACGCGCAGGACGAGGTGGCAACAAATTCCTTCGGCCAAATGACGAGCCGAG 3661
Db 9879 CCAATGCAACGCGCAGGACGAGGTGGCAACAAATTCCTTCGGCCAAATGACGAGCCGAG 9938
Qy 3662 TTACAGAGCTCATTTAGCATTTTCCCAAGGAGGCGGCGAGGCGCGGTTGGT 3721
Db 9939 TTACAGAGCTCATTTAGCATTTTCCCAAGGAGGCGGCGAGGCGCGGTTGGT 9998
Qy 3722 GTGTGTGCGGTGTCGGCAGCATCCCGCGCGCTGCTGCGGTGCGCGGAGCTCGGCT 3781
Db 9999 GTGTGTGCGGTGTCGGCAGCATCCCGCGCGCTGCTGCGGTGCGCGGAGCTCGGCT 10058
Qy 3782 CTGTCTCTCTCCCTCCCGCTTACCTTCCACGCGGAGCCGCGCGCAGTCAACTCT 3841
Db 10059 CTGTCTCTCTCCCTCCCGCTTACCTTCCACGCGGAGCCGCGCGCAGTCAACTCT 10118
Qy 3842 CGCATTTGCGCTGCTTGGCAGCGGATAAAGGGGCTGAGGAATACCGGACACGCTC 3901
Db 10119 CGCATTTGCGCTGCTTGGCAGCGGATAAAGGGGCTGAGGAATACCGGACACGCTC 10178
Qy 3902 ACCCGTTGC 3910
Db 10179 ACCCGTTGC 10187

RESULT 3
US-10-027-632-24494/c
; Sequence 24494, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24494
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24494

Query Match 16.1%; Score 630.6; DB 13; Length 631;
Best Local Similarity 99.8%; Pred. No. 2.1e-158;
Matches 630; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 968 AAGGCTCATATGTTGCGATGATCCCGCAGCTCCACCTTCCAAAGTCTCATCTTATC 1027

Db 631 AAGTCTCCTATGTTGCCATGATCCCGCCACCTCCACTCCCAAGGCTCATCTTATC 572
QY 1028 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTTCACCATTAATGAATAA 1087
Db 571 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTTCACCATTAATGAATAA 512
QY 1088 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 452
QY 1148 TTATATCTAGGAGTGGAAATGCTGATCCCGGTAATATTTTGACAGGCAGAGTTTCAGGG 1207
Db 451 TTATATCTAGGAGTGGAAATGCTGATCCCGGTAATATTTTGACAGGCAGAGTTTCAGGG 392
QY 1208 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTC 332
QY 1268 GGAGTTTATTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTATACT 1327
Db 331 GGAGTTTATTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTATACT 272
QY 1328 AAGTCAGAGTGCAGAAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATCAAGCATTT 1387
Db 271 AAGTCAGAGTGCAGAAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATCAAGCATTT 212
QY 1388 TGAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATTAATAAATCAAAATTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATTAATAAATCAAAATTTTTT 152
QY 1448 TTTTTCATTTGAGAAAGAGTCTTCTCTTTTACCCCTGGCTGGAGGCGAGTGTGATCTC 1507
Db 151 TTTTTCATTTGAGAAAGAGTCTTCTCTTTTACCCCTGGCTGGAGGCGAGTGTGATCTC 92
QY 1508 AGCTCACTGCAACTTTCCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 1567
Db 91 AGCTCACTGCAACTTTCCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 32
QY 1568 AGCTGGGACTACAGGCACCTTCCCAACCATGCC 1598
Db 31 AGCTGGGACTACAGGCACCTTCCCAACCATGCC 1

RESULT 4
US-10-027-632-24494/c
; Sequence 24494, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24494
; LENGTH: 631
; TYPE: DNA

; ORGANISM: Human
US-10-027-632-24494
Query Match 16.1%; Score 630.6; DB 16; Length 631;
Best Local Similarity 99.8%; Pred. No. 2.1e-158;
Matches 630; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 968 AAGTCTCCTATGTTGCCATGATCCCGCCACCTCCACTCCCAAGGCTCATCTTATC 1027
Db 631 AAGTCTCCTATGTTGCCATGATCCCGCCACCTCCACTCCCAAGGCTCATCTTATC 572
QY 1028 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTTCACCATTAATGAATAA 1087
Db 571 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTTCACCATTAATGAATAA 512
QY 1088 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATCATGTATATCATTTGTGTGGGCATATGTTTTCATTCTGTGGGT 452
QY 1148 TTATATCTAGGAGTGGAAATGCTGATCCCGGTAATATTTTGACAGGCAGAGTTTCAGGG 1207
Db 451 TTATATCTAGGAGTGGAAATGCTGATCCCGGTAATATTTTGACAGGCAGAGTTTCAGGG 392
QY 1208 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTC 1267
Db 391 GAAGAAAAAATTTGGGAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTC 332
QY 1268 GGAGTTTATTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTATACT 1327
Db 331 GGAGTTTATTTATATCTGTTGACAAATGTCAGTTTGATGAAGATACAAAGTTATACT 272
QY 1328 AAGTCAGAGTGCAGAAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATCAAGCATTT 1387
Db 271 AAGTCAGAGTGCAGAAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATCAAGCATTT 212
QY 1388 TGAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATTAATAAATCAAAATTTTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGCTGTAAACAAATAATTAATAAATCAAAATTTTTT 152
QY 1448 TTTTTCATTTGAGAAAGAGTCTTCTCTTTTACCCCTGGCTGGAGGCGAGTGTGATCTC 1507
Db 151 TTTTTCATTTGAGAAAGAGTCTTCTCTTTTACCCCTGGCTGGAGGCGAGTGTGATCTC 92
QY 1508 AGCTCACTGCAACTTTCCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 1567
Db 91 AGCTCACTGCAACTTTCCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGT 32
QY 1568 AGCTGGGACTACAGGCACCTTCCCAACCATGCC 1598
Db 31 AGCTGGGACTACAGGCACCTTCCCAACCATGCC 1

RESULT 5
US-10-087-192-814/c
; Sequence 814, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 814
; LENGTH: 59725
; TYPE: DNA

[illegible]

```

; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1569
; LENGTH: 9631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(9631)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(230)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4089)..(4448)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9336)..(9431)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (3602)..(3701)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (9233)..(9332)
; OTHER INFORMATION: a, t, c, g, unknown or other
; US-10-017-161-1569

Query Match 10.4%; Score 407.4; DB 15; Length 9631;
Best Local Similarity 58.6%; Pred. No. 1.2e-97;
Matches 1078; Conservative 0; Mismatches 631; Indels 132; Gaps 16;

Qy 358 TAATCCATATCATATTATAGACAGAGTTCTCAGCTGTCACCTGTCACCCAGGCTGGA-GTGCAG 416
Db 274 TATTATTATTATTATTTTGTAGACAGGCGCTTGTCTGTGACCCAGGCTGGAGGTGCAG 333
Qy 417 TGGCACAATCATAGCTCAGCTGTAACTCAAACCTCTGGGCTCAAGTGATCTACTACCTC 476
Db 334 TAGCATGATCATAGCTCAGCTGCAGCTTCAACCTCTGGGCTCAATGATCTTCCACCTC 393
Qy 477 AGCTTCAGAGTAGTAGGACTAGACGACACACAGCAGCATACCTGGCTAAATTTTTTTT 536
Db 394 AGCTTCCTAGTAGTGCAGACTGCGAGCTGAGCATGCACCACAGCCCTGGCTTAATTTTGT 452
Qy 537 TAAATTTTCAATTTATGTATTCATTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
Db 453 -----GTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
Qy 597 TCTCAGTTTGTATCCAGGCTGAGGCGAGTGGCATGGTGACAGCTGAGCAGCCTTCACT 656
Db 482 TTTTACCGTGTGCCAGGCTG-----GTCCCTGAAC 512
Qy 657 TCCTGGGCTCAAGTGATCCTCTCGCTCAGCTCCCAAGTAGCTGGGACTACAAACAGT 716
Db 513 TCCTGGGCTCAAGCCATCTGCCAGCTTCGTCTTCCAAAGTGTGGGATACAGGTATAA 572
Qy 717 GTCACCATGCTGGCTGAT-----ATTTTTTTTCTTGAAAACAGGGTATCATCTCT 765
Db 573 GCCACCAACAGGCGCTATTTTAATTTTATTTTACTTTTTTGTAGTCAGGCTCTTGTGT 632
Qy 766 GTTGCCCGAGCTGGAGTACAGTGGCGTAATAATAGCTCACTGCAGCTC-CCCTCCCTGGG 824
Db 633 GTCACCCAGGCTGGAGTGCAGTGGCAAAATACAGCTCACTGCAGCTCAACCTTCCTGGG 692
Qy 825 CTCAGCAATCCGCTGGCTCAGCATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACCA 884
Db 693 CTCAGTGAATCCTCTCGCTCAGCTGCCATAATAGCTAGGAGCAGAGGTGTGTGCCACAA 752
Qy 885 GGCCCGACTAAGTTTTAAAAAATGATTTTTTGGTATAGAGAGGTCTTGTGTATGTGTCTCA 944

```

573	GCACCAACGGCCATTTTAATTTTATTTTACTTTTTTGATTCAGGCTCTTGCTGT	633
766	GTTGCCCAGGCTGGAGTACAGTGGCGTAAATAATAGCTCACTGCAGCCTC-CCTCCCTGGG	824
633	GTCAACCAGGCTGGAGTGCAGTGCACAATATACAGTCTCACTGCAGCCTCAACCTCTCTGGG	692
825	CTCAAGCAATCCGTGCTGCCCTCAGCATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACCA	884
693	CTCAAGTGATCCTCCTGCCCTCAGCTGCCCTTAATAGCTTAGGAGCACAGGTGTGTGCCAAA	752
885	GGCCCAAGCTAAGTTTTTAAAAAATGATTTTTTGGTAAATAGAGGAGGTCTTGTCTAATGTTGTCA	944

QY	1860	CAGGCGTGGTGGCTCATGCTGTAAATCCAGACACTTTGGAGGCCAAGGCATGTGGATCA	1919
DB	10583	CAGGCATGTGGCTCATGCTGTAAATCCAGACACTTTGGAGGGTGAGCGAGGCGATCA	10524
QY	1920	CTGAGGTGAGGATTTCAAAAACAGCTGGCCAAACATGGTGAACCCCGTCTCTACTAAA	1979
DB	10523	CCTGAGGTGAGGATTTGAGATCAGCCTGGCCAAACATGGTGAACCCCTGCTCTACTAAA	10464
QY	1980	AATACAAAAATTAGCTAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTTGGGAGGG	2039
DB	10463	AATAC-AAAAATTAACCGGCGTGGTGGCAGGCGCTGTAAATCCAGGTACTTCGGGAGGC	10405
QY	2040	TGAGGCAGGAGAAATTTCTTGAACCCGGGAGGCAGAGGTTGCAGTGAAGCGAGATCAACAC	2099
DB	10404	TGAGGCAGGAGATCGCTTGTACTTGGAAAGGCAGAGGTTGCAGTGAAGCGACACTGAACC	10345
QY	2100	ACTGCACCTCAGCCTGGGGGAGAGCGAGACTTCTCTCAAAAAAACAACAAACAAAG	2159
DB	10344	ATTGCACTCAGCCTGGGTGACAGAGCAAGACTCTGTGCGCAAAAAAATAAATTATA	10285
QY	2160	AATTAA 2165	
DB	10284	ATTTTA 10279	
RESULT 9			
US-10-240-425-1584/C			
; Sequence 1584, Application US/10240425			
; Publication No. US20040033502A1			
; GENERAL INFORMATION:			
; APPLICANT: Williams, Amanda			
; APPLICANT: Bolland, Joseph F.			
; APPLICANT: Lord, Reginald V.			
; APPLICANT: Alvarez, Chris			
; APPLICANT: Wetzel, Jon C.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Vockley, Joseph G.			
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue			
; FILE REFERENCE: 44921-5026			
; CURRENT APPLICATION NUMBER: US/10/240,425			
; CURRENT FILING DATE: 2002-09-30			
; PRIOR APPLICATION NUMBER: PCT/US01/09847			
; PRIOR FILING DATE: 2001-03-28			
; PRIOR APPLICATION NUMBER: US 60/193,446			
; PRIOR FILING DATE: 2000-03-31			
; NUMBER OF SEQ ID NOS: 1588			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1584			
; LENGTH: 36296			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 282180			
US-10-240-425-1584			
Query Match 9.0%; Score 353.6; DB 13; Length 36296;			
Best Local Similarity 57.2%; Pred. No. 8.4e-83; Indels 107; Gaps 13;			
Matches 956; Conservative 0; Mismatches 609;			
QY	522	GCTAATTTTTTTTTTTTTTAAATTTTCATTTATGTATTCAATTTCTTTCTTTTGTGTGT	581
DB	15559	GGTAATGTTGAGCTTTCTCAAGCAGCATATTATCTATTATTATTTATTCATGATTATT	15500
QY	582	TGTTTTGAGATAGGTTCTCATCTTTGTTACCCAGGCTGGAGGGCAGTGGCATGGTGACAGC	641
DB	15499	TTTTTTAGACACAAGGTTCTCATTCCTCCACCGGCTGGAGTAGTTGTGCAATCATAGCTC	15440
QY	642	TGACAGCCTTGACTTCTCGGCTCAAGTGATCCTCTCGCTCAGCCTCCCAAGTAGCTG	701
DB	15439	ACTGCAGCCTCCAACTCCTGGGCTCAAGTGATCCTCACACCTCAGCCTCCTGAGTAGTTG	15380
QY	702	GGACTCAAAACAGCTGTGCACATCGCTGGCTGATATTTTTTTCTTTGAACACAGGATCA	761

15379	DB	GGACATGCAAGTGCATGCCACCAAGCTGGCCCTTTTTTTTTTTTTCAGACAGATTTTCGCT	15320
762	QY	CTCTGTTGCCCAGGCTGGAGTACAGTGGCGTAATAATAGCTCAGCTGCAGCTCC-CCTCC	820
15319	DB	CT-TGTACCCCAAGCTGGAGTGCAGTGGCAGATCTTGCTCAGCTGCACCTCCGCTCC	15261
821	QY	TGGGCTCAAGCAATCCGCTGGCCTCAGCATCTCTGATGCTGGGACTACAGCTTTGTGCC	880
15260	DB	CAGGTTCAAGTGATTCTCTGCGCTCAGCCTCTCGAATAGCTGGGATTACAGGCAACCAACC	15201
881	QY	ACGAGGCCAGCTAGTTTAAAAAATGATTTTTGGTATAGAGGAGGCTTTGCTATGTTG	940
15200	DB	ACCAAGCCTGTCTAAATTTTT-----TGTATTTTATGAGACAGGGTTTACCAATGTTG	15146
941	QY	CTCAGGCTGTATTTTATTTGTTGACAGAAGGCTCACTATGTTGCCATGATCCCCCAACC	1000
15145	DB	GGCAAGCTGCTCTAACTCTCGACCTCAG-----TGATCCGCTGCC	15102
1001	QY	TCCACTTCCCAAAGTCTCATCTTATCTGTTCAATTAGTCAGTTGACAGCATTTAGTTG	1067
15101	DB	TCAAGCCTCCCAAAGTCTGGGATTACAGGTGTGAGGGCCATTT--TTTTTTTAAAGGTAC	15044
1061	QY	TTTTCACTTTTTGACCATTAATGAATTAATCTCCAGTGAATATTCATGTATATACATTTGT	1120
15043	DB	AGATGGGGCTTGTCTATGTTGCTCAGGCTGGTCTCAAAATCTTTGGGCTCAAGTGATCTCTC	14984
1121	QY	GGGCATATGTTTTCATTTCTGTTGGTTTATATCTAGGAGTGGAAATTCGTGATCCCGGG	1180
14983	DB	CCTCTGAGCCTCCCAAAGTGTGGGATTTAGGCAGGAGCACCGCACCCAGCCTCAGG	14924
1181	QY	TAATATTTTGACAGGCAGAGTTTCAGGGGAAGAAAACTTGGGAAATGAAGCATGTTTAG	1240
14923	DB	TGACATTTTAAATGGGAGAGGATGGTGGCTGGAGTCAATCAAGGGGACACGCTCCTGATG	14864
1241	QY	AAATCAGCAGAGTGCAGGGGTTTTTCGGAGTTTTATTTTATTTCTGTTGACAAATGTTG	1300
14863	DB	CTGTTAGGAGCCAGCTAGAGTTTGGTCTTCCCTTAGTCTGAGTTAGGAGAGAGTTT	14804
1301	QY	CAGTTTGATGAAGATCAACAGTTTATACTAAGTCAGAAGTGAGAAATTAAGGCTGGAATAGGG	1360
14803	DB	CTGTGCCAAGAGCCCTGCAAGTT-----	14782
1361	QY	CGTTCAGAGTAAATCATGAAGCACTTTGAATACCAAAATTAAGAGGCTTGCTGTAAAC	1420
14781	DB	-----CAGTGGTGGGACTGACACTCCTCTTATAGGAAGACACC	14741
1421	QY	AAAATTAATAAAATCACAATTTTTTTTTTTTTTTTGGAAAGAGCTTGCTCTTTTCCACC	1480
14740	DB	CTCTCTGCTAGTCTTTTTTTTTTTTTTTTTTTTGGACAAGAGCTTGTGTTCTTTCA-C	14682
1481	QY	CTGGCTGGAGGGCAGTGTGATCTCAGCTCAGCTGCAACTTTGCGCTCCCGGGTTCAAG	1540
14681	DB	CAGGCTGGAGTGCAATGGCATGATCTGGCTCAGCTGCAACTCCACCTCTTGGGTTCAAG	14622
1541	QY	CAATCTCCTGCTTCAGCCTCCCAAGTAGCTGGGACTACAGGCACTTCCCAATGCCCCA	1600
14621	DB	CGATTTCTTCTGCTCAGCCTCCGAGTAGCTGGGAGTACAGGCAACCCGCGACCAATCTG	14562
1601	QY	GCTGATTTTTTGATTTTTTAGTAGAGATGGGATTTCACTTTGTTGGCCAGCTGTTCTCAA	1660
14561	DB	GCTAAATTTTGTATTTTAAATAGAGGGGAGTTTCAACCAATGTTGGCCAGATGGTCTCGA	14502
1661	QY	ACTTTTTGCTGTCAATTTGTTGAATTTGTTTCC-----TTTTGCTGAGGTAGGGC	1713
14501	DB	TTCTCTGACCTATGATCTGCTGGCTCCGGTTTACAGGCATGAGCCACTGCGCCCGGCC	14442
1714	QY	CCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGGTTTGACCACTGTC	1773
14441	DB	CCTCTTGCTATTCTTCAGCAGCCTTGGAGTGAGAAACCTGAGGCCCTGCTGCTTCCCTGG	14382
1774	QY	ACTTGAAGACCAAGTGTGACCAAGGCGCTTC-AGGAGTAGAGTGATCTCTGCTCGAAGA	1832

[illegible]

Db	4589	GTCCACCATGCCCACTTAATTTTTTTTTTGAGATGGGGTC-----	4622
Qy	568	TTTTTTTGTTGTTGTTTGGATAGGGTCTCACATTTGTTTACCCAGGCTGGAGGCAGT	627
Db	4629	-----GTTCTGCTCTGTCACCCAGGCTTGGAGTGCAGT	4660
Qy	628	GGCATGGTGA CAGCTGA--GCAGCCCTTGACTCTCTGGGCTCAAGTGATCTCTCGCTCA	685
Db	4661	GGCATGATCTCCGCTCACTGCTCCCTGCTCTCTGGGCTCAAGGGATCTCCCACTCA	4720
Qy	686	GCCTCCCAAGTAGCTGGGACTACAAACAGTGTCACCATGCTGGCTGGCTGATATTTTTTTC	745
Db	4721	GCCTCCCAAGTAGCTGGGCTACAGGCACAAGCATGGAGCCCAAGCTAATTTTTTGTATT	4780
Qy	746	TT-----GAAACAGGGTATCACATCTGTGTGCCAGGCTGGAGTACAGTGGCGTAATAATAG	800
Db	4781	TTTGGGCACAGCAGGGTTTCCATTTTGGCCAGGCTGGTCTCAA-----	4825
Qy	801	CTCACTGACGCTCCCTCTCTGGGCTCAAGCAATCCGCTGGGCTCAGCATCTCTGAGTAGC	860
Db	4826	-----ACTCTGAGCTCACTGATCTTCCACCTGGGCTCAAAAGTGC	4870
Qy	861	TGGGACTACAGGCTTGGCCACAGGCCACAGCTAAAGTTTTTAAAAAATTTTTTGGTATA	920
Db	4871	CAGGATTCAGGTTTGAGCCACCGCACCTGGCCTAATTTTTGTATTTTTTTTTTTTTTT	4930
Qy	921	GAGGAGTCTTGCTATGTGCTCAGGCTGTATTTTTTATTGTTGAGACAAGGCTCACTAT	980
Db	4931	TTTTGTAGATGGGGTTGGCCATGTGGCCAGGCTGGTAAACAGCAATTTTTCTTCT	4990
Qy	981	GTTGCCATGATCCCCCACTCCACTTCCCAAAGTGCTCATCTTATCTGTTCATTAGTCA	1040
Db	4991	TCCTCTTTTTTCTGAGATGGAGTTTAGCTCTTGGAGTGTGATGGCATGATCTTGGCTCA	5050
Qy	1041	GTTGACAGCAATTTAGTTGTTTCACATTTTTCACCATTTAGCAATTAATCTCCAGTGAAT	1100
Db	5051	CTCAACCTCTGCTCTCTGGGTTTCAAGCAATTTCTCC---TGCCCTCATCTTCCCAAGTAG	5106
Qy	1101	ATTCAATGATATATTTGTTGGGCATATGTTTTCAATTTCTGTTGGGTTTTATCTAGGAG	1160
Db	5107	CTGGGATACAGGTGCTGCCCCACACCAGCTAATTTTTTGTCTTTAGTAGATGG	5166
Qy	1161	TGGAATTGCTGGATCCGGGTAAATATTTTACAGGCAGAGTTTCAAGGGA-----AGAA	1213
Db	5167	GGCTTCATCATATTTGGCCAGATGGTCTTGAATCTCTGACCTCAGGTGATCCGCCATCT	5226
Qy	1214	AAACTTGGGAAATGAAGCATGTTTGAATATCAGCAAGAGTGCAGGGGTTTTTCGAGTT	1273
Db	5227	CAGCCTCCCAAGTGTGGGATTTATAGTGTGAGCCACCATGCTGGCCTTTTCTTCTTT	5286
Qy	1274	TTATTTTATATCTGTTGACAAATGT---GCAGTTTGTATGAAGATACAAGTTATATAAG	1330
Db	5287	TCTTTTTTTTTTTTTTTAGACAGGTTCTCGCTTGTCACCAGACTGGAGTGCAGTGGCA	5346
Qy	1331	TGAAAGTGAGATTAAGGCTTGAATAGGCGGTTTCAAGATAAAATCATGAAGCACTTGA	1390
Db	5347	CCATCTCGGCTCATTAACAACCTCCACCTCCAGTATCAAGTAAATCTCGTGCCTCAGCCT	5406
Qy	1391	ATACCAAAATTAGGAGCTTGGTGTAAACAAATATATAAAAATCAACAATTTTTTTTTT	1450
Db	5407	CTCCAGTAGCTGGGACTACAGGTGCCACGCCAGGCTCTGGTCTGGCTTTTTTTTTTT	5466
Qy	1451	TTTTTTTTGAGAAAGTCTTGTCTTTTACCTCGGCTGGAGGGCAGTGTGTGATCTCAGC	1510
Db	5467	TTTTTTTTGAGACAGAGTCTCGCTGTTGCCAGGCTGGAGTGCAGTGTGTGATGATAGC	5526
Qy	1511	TCACTGCAATTTTCGCTCTCCGGGTTTCAAGCAATTTCTCTGCTTCAGCCTCCCAAGTAGC	1570
Db	5527	TCACTGCGCTCCAATCTCTGACTCAAGCAATCTCCAGCCTCAGCCTCCCGAATAGC	5586
Qy	1571	TGGGACTACAGGCATTTCCCAACATGCCCAAGCTGATTTTTTGTATTTTTTATGATAGATGG	1630
Db	5587	TGGGATTTATAGGCATGAGGCACCATGCCCAAGCTAATTTCTGTATTTTTTATGACACAGG	5646

Db 4214 GAGGACAGAGAAATCACTTGAACCCAGGAGCGAGGTTGCAGTGCAGCCAGATCACACCA 4273
QY 2101 CTGCACTCCAGCTCGGGGAGAGAGAGAGACTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 4274 CTGCACTCCAGCTCGGGGACAGAGTGGCTGTGCTCTCAAAAAACAAAAACAAAA 4333
QY 2161 ATTAAGCAAAATTAGACA 2177
Db 4334 ACAAAACAAAAACAAAA 4350

RESULT 13
US-09-764-891-7809
; Sequence 7809, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7809
; LENGTH: 29163
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7809

Query Match 8.9%; Score 347.4; DB 10; Length 29163;
Best Local Similarity 69.9%; Pred. No. 3.4e-81;
Matches 515; Conservative 0; Mismatches 211; Indels 11; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGTGAGAAAGAGTCTGTCTTTCAACCTGGCTGGCTGGAGGCGAGTGGTG 1500
Db 3623 TCTTTCTTTTCTTTTCTGAGTCTTGTCTTGTGCGCCAGGCTGGAGTGAGTGGTG 3682

QY 1501 TGAATCAGTCACTGCAACTTTTGGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCCT 1560
Db 3683 CGATCTCAGTCACTGCAAACTCCACTCCCGGGTTCAAGAGATTTCTCTGCTTAGCCT 3742

QY 1561 CCCAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTAG 1620
Db 3743 CTTGAGTAGCTGGGATACAGGCTGCACTACCGTGCCTGGCTAAATTTATATTTTAA 3802

QY 1621 TAGAGATGGGATTTCACTTTTGTGGCCAAAGCTGGTCTCAAACTTTTGTGTCATATTTG 1680
Db 3803 TACAGATGGGTTTCAACCATTTTGGTTCAGGCTGGTCTCGAACTCTGACTTCATGATCCG 3862

QY 1681 TTGTAACTATTTGTTCTTTGCTGAGGTAGGGCCCCCAGACCAAAAAATAATCTTAG 1740
Db 3863 TCCCCCTCGACCTCCCAAAAGTTAGGATTTACAGGCATGAGCCACTGTGCCAGCC---- 3918

QY 1741 AATCCAAATCAGTGTGTGTTTGTGACACTGTACATTTGAGAACACACAGTGTGACCAAGGC 1800
Db 3919 ---GCTAAATTTTGTACAGACAGAGTTTCTCCATGTTGCCAGGCTGATCTGGAA 3975

QY 1801 CTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATAATTTCCGGGCC 1860
Db 3976 CTCCTGGGCCAAGTGAACCATCTGCTCGC---CTCCAAAGTGTGGGATTTACAGGCC 4032

QY 1861 AGCGTGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCTGTGGATCAC 1920
Db 4033 GGGTGGCGGTGGCTCAGCGCTGTAATTTCCAGCACTTTGGGAGGCCAAGGCTGGATCAG 4092

QY 1921 CTCAGGTCAGGAGTTCAAAACAGCTGGCCAACTGGTGAACCCCGTCTCTACTAAAA 1980
Db 4093 CTGAGGTCCGGAGTTCAAGACAGCTGGCCAAAGGCTGACCCCTGCTCTACTAAAA 4152

QY 1981 ATACAAAAATTAGCTAAGTGTGGTGGCGCATGCCTGTAAATCCAGCTACTTGGGAGGT 2040
Db 4153 ATAC-AAAAATAGCCAGGTGTGGTGGCGCATGCCTATATATATATATATATATATAT 4211

QY 2041 GAGGACAGAGAAATTTCTTTGAAACCCGAGGAGCGAGGTTGCAGTGAAGCGAGATCACACCA 2100
Db 4212 GAGGACAGAGAAATCACTTTGAACCCAGGAGCGGAGGTTGCAGTGAAGCGATCACACCA 4271
QY 2101 CTGCACTCCAGCTCGGGGAGAGAGAGACTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 4272 CTGCACTCCAGCTCGGGGACAGAGTGGCTGTGCTCTCAAAAAACAAAAACAAAA 4331
QY 2161 ATTAAGCAAAATTAGACA 2177
Db 4332 ACAAAACAAAAACAAAA 4348

RESULT 14
US-09-880-107-2278
; Sequence 2278, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Logic, Inc.
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2278
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M17262
US-09-880-107-2278

Query Match 8.8%; Score 342.8; DB 9; Length 26928;
Best Local Similarity 57.5%; Pred. No. 5.5e-80;
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

QY 735 TATTTTCTTTTCTTGAACAGGGTATCATCTCTGTCGCCAGGCTGGAGTACAGTGGGTAA 794
Db 3122 TTTTCTTTTGGGAGTCAAGAGTCTGTCTCTGTCGCCAGGCTGGAGTCAAGTACCA 3181

QY 795 TAATAGCTCACTGCAGCCCTCC--CCTCCTGGGCTCAAGCAATCCGCTGGCCTCAGCATCT 853
Db 3182 TCTAGCTCACTGCAGCCCTCCGCTCCCGGTTCAAGCTATTCTCTGTCTCAGCCTCC 3241

QY 854 GAGTAGCTGGGACTACAGCTTGTGCCACAGCCAGCTAAGTTTAAATAATGATTTT 913
Db 3242 AAATAGCTGAGACTATAGGCACGCATCTCATGCCAGCTAATTTT-----TTATTTT 3295

QY 914 TGGTATAGGAGGCTTGTGCTATGTTGCTCAGGCTGATTTTATTTTGTGAGCAAGGTC 973
Db 3296 TAGTAGAGACGAGGTGTCTCCATGTTGGCCAGGTTGGTCTTGAAC----- 3341

QY 974 TCACATATGTTGCATGATCCCGCCACCTCCACTTCCAAAGTGCATCTTATCTGTTC 1033
Db 3342 ---CCTGTCTCAAGTATCCACTGCTCGGCTCCCAAGTGGGATTCAGGCAATG 3399

QY 1034 TTAGTCAGTTGACAGACATTTAGTGTGTTTCCACTTTTGGACCATATGAAATACTCC 1093
Db 3400 AGACACCGCGCCCGGCC-----TGCCTTGTCCCTTCTTAAATAGTGTCTCCATTTGTA 3454

QY 1094 AGTGAATATTCATGATATATATGTTGTTGGCATATGTTTCTATTTCTGTTGGTTTAT 1153
Db 3455 GCTGCTGATTTCTTTGGGACATTTGCTCCGTAACCTTTTCATAAGCATCAGTATTTCA 3514

Qy	1154	CTAGGAGTGGAAATTGCTGGATCCCGGGTAAATATTTGACAGGCAGAGTTTCAGGGGAGAA	1213
Db	3515	CCATTCTTCACCAACCAAGCTTCACCGTAAATTTGTTGTTCTGCTTCAATTTTCAGCA	3574
Qy	1214	AAACTTTGGGAAAATGAAGCATGTTTAGAAATCAGCAAGAGTGCAGGGGTTTTTCGGAGTT	1273
Db	3575	GAATTCAATTAGCTCTGTATGAAGGCTCGCTTCAAACTGATGCTATCCTTCTTAGTGCC	3634
Qy	1274	TTATTTTATATTCGTGTGACAAATGTGC-----AGTTTGATGAAGATACAAAGTTATCTA	1328
Db	3635	TCAAACTACATCTCTGTCTCACTCATGTTATAGCAAGTTAGTGTGAGTTATTTTGGTGCAC	3694
Qy	1329	AGTGAGAGAGTGAAGATTAAGGCTGGBATAGGGCGTTCAGAGTAAGATCATGAGCACTTTT	1388
Db	3695	AAAAATTTTTTAAATCCATGCSAGTCTTTTTTCATAATAGCAATTTTCCATGAACTTTTC	3754
Qy	1389	GAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAAATTTTTTTTT	1448
Db	3755	GAAGACCCCTTGTAGATGCTGTGTGTTTAAACCAACCCAGTTTA-----CAGTAATTTTTT	3810
Qy	1449	TTTTTTTTTGAAGAAGAGTCTGTCTCTTTTACCCTTGGCTGGAGGCGAGTGGTGTGATCTCA	1508
Db	3811	TTTTTTTTTGAATGAAGTCTTGTCTCTGTGCCCAAGGCTGGAGTGCATTTGGCAACACTCTG	3870
Qy	1509	GCTCACTGCACATTTCCGCTCCCGGGTTCAAGCAATTTCTCCTGCTTCAGGCTCCCAAGTA	1568
Db	3871	GCTCAGTCGAACCTCTCGCTCTCGGGTTCAAGCAATTTTCTCTCAGTCTCCCGAGTA	3930
Qy	1569	GCTGGGACTACAGGCACATTCGCCACATGCCACGCTGATTTTTGTATTTTATGATAGATG	1628
Db	3931	GCTGGGATTAACAGGTGTGTGCCACCATGCTAGCTAATTTATGTGTTTTTATGATAGAGCG	3990
Qy	1629	GGATTTTCACTTTGTGGCCAGCTGTGCTCAAACTTTTTCGCTGTCATATTTGTTGTAAC	1688
Db	3991	GGGTTTTCACATGTTGGCTAGCTGTGCTCGAACTCTCACTTTGTGATCGGCCCGGCTC	4050
Qy	1689	ATTGTTCCCTTTTGTGTAGTAGTGGGCCCCAGACCAAAAAATAATCTTTAGAATCCAAA	1748
Db	4051	GGCCTCCCAAGATATTGGGATTACAGCGCTGA-----GACTCTTGCACTTGGCC	4099
Qy	1749	TCAGTGTGTTGGTTTGAACACTGTCTCTTGAGAACCAAGTGTGACCAAGGCTCTCAGGAG	1808
Db	4100	TACAGTAATTTTATAGCAGCGCTAGGCTTAAGATAGCCATTTCTGGGTATAGAATGTCA	4159
Qy	1809	TAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATAATTTCTCGGSCCAGGCGGTGG	1868
Db	4160	TACTGAACAGCCCTTGAAGTGTGAGTAAAGTCTGCAAGAGGCC-----GGGCAG	4210
Qy	1869	TGGCTCATGCTGTAAATCCAGCACTTTGGAGGCCCAAGCGATGTGATCACTGAGGTC	1928
Db	4211	TGGCTCATACCTGTAAATCCAGCACTTTGGGGGCCGAGGCGATGATCACTGAGGTC	4270
Qy	1929	AGAGTTTCAAAAACAGCGCTCGCCAACTGTGAAACCCGCTCTCTACTTAAAAATACAAAA	1988
Db	4271	AGCAGTTGAGACCAAGCGCTGACCAACATGTGTGAACCCCACTCTACTAAAAATAC--AA	4328
Qy	1989	AATTAGCTAAGTGTGGTGGCGCATGCTCTGTAATCCC-----AGCTACTTG	2033
Db	4329	AATTAGCTGGSCGTGGTAGTGCATGCTTGTAAATCCCTAGCATGCACTTGGGAGCTACTTG	4388
Qy	2034	GGAGGTGAGCAGGAGAAATTTCTTGAACCCGGAGGCGAGGTTGCAGTGGAGCGAGAT	2093
Db	4389	GGAGGCTGAGCAGGAGAAATCACTTTGTACTCAGAGGCCGAGGTTGCAGTGTGAGCTGAG	4448
Qy	2094	CACACCACTGCACCTCCAGCCTTGGGGGAGAGACGAGACTTCTCTCTCAAAAAAACA	2153
Db	4449	CACGCCACTGCATCTCTTTCTGGGTGACAGATGAGACTCCATCTCAAAAAACA	4508
Qy	2154	CAAAAGAAATTAAGCAATTAAGCATTTGCAGAGAG	2187
Db	4509	AAACAAAAACAAAAACAAACAAAAAACCCACAG	4542

RESULT 15
US-10-235-192A-34
; Sequence 34, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: Methods and Compositions for Identifying
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases
; TITLE OF INVENTION: and Disorders Associated Therewith
; FILE REFERENCE: WMI-011
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-235-192A-34

Query Match 8.8%; Score 342.8; DB 13; Length 26928;
Best Local Similarity 57.5%; Pred. No. 5.5e-80;
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

QY 735 TATTTTTTTTCTTTGAACACGGGTATCACTCTGTTTGCCCGAGGCTGGAGTACAGTGGCGTAA 794
Db 3122 TTTTTTTTTTGGGAGTCAAGAGCTTGCTCTGTTTGCCCGAGGCTGGAGTGGAGTGCACGA 3181

QY 795 TAATAGCTCACTGCAGGCTCC-CCTCCTGGGCTCAGCAATCCGCTGGCTCAGCATCCT 853
Db 3182 TCTCAGCTCACTGCAGGCTCCGGCTCCGGGTTCAAGCTAATCTCTCTGCTCAGCCTCC 3241

QY 854 GAGTAGCTGGGACTACAGGCTTGTCGCCACGAGGCCCACTAAGTTTTTAAAAATGATTTT 913
Db 3242 AAATAGCTGAGACTATAGGCACGCACATCCATGCCAGCTAAATTTT-TTATTTT 3295

QY 914 TGGTATAGAGAGGCTTGCTATGTTGCTCAGGCTGTAATTTTATTTGTTGAGCAAGTTC 973
Db 3296 TAGTAGAGACGAGGTGTCTCATGGTGGCAGGTGGTCTTTGAAC-TTTTTT----- 3341

QY 974 TCACATAGTTGCCATGATCCCCACCCTCCACTTCCCAAAGTGCTCATCTTATCTGTTC 1033
Db 3342 --CCTGTCTCAAGTGAATCACTGCGCTCCGCGCTCCCAAAGTGGTGGATTCGAGCATG 3399

QY 1034 TTAGTCAGTTGACAGACATTTAGGTTGTTTCCATTTTTTGACCATATAGTAATACTCC 1093
Db 3400 AGACACGCGCCGCGC-----TGCTTGTGCTCTTAAATAGAGTTGCTCATTTGTAA 3454

QY 1094 AGTGAATTTCAATGATACATTTGTGGGCATATGTTTTCAATTCGTGGGTTTATAT 1153
Db 3455 GCTCTGATTTCTTTGGGACATGCTCCGTAACATCTTTCAATAGCATCAGTCAATTC 3514

QY 1154 CTAGGAGTGGAAATGCTGGATCCCGGGTAATATTTTGCAGGCGAGGTTCAGGGGAAGAA 1213
Db 3515 CCAATCTTCCACCCCAAGCTTCCACGTAATTTTGTTGTTGCTCTTGTCTCAATTTTCAGCA 3574

QY 1214 AAACCTGGGAAATGAAGCATGTTAGAAATCAGCAGAGTGCAGGGGTTTTTCGGAGTT 1273
Db 3575 GAATTCATTTAGCTCTGATAGGCGCTCCGCTTCAAACTGATGTCTTATCCCTCTTAGTGCC 3634

QY 1274 TTATTTTATATTCGTGTGACAAATGTGC-----AGTTTGATGAAGATACAAGTTTACTA 1328
Db 3635 TCAAACTTACATCTCTGTTCACTCATGTTATAGCAAGTTAGTGTGAGTTTATTTGGTGCAC 3694

QY 1329 AGTCAGAGTGAAGATTAAGGCTCGAATAGGCGGTTTCCAGAGTAAATCATGAAGCATTT 1388
Db 3695 AAAAAATTTTTTAAATCCAATGAGTCTTTTTTTCATAATAGCATTTTCCATGAACTTTTC 3754

QY 1389 GAATACCAAATTAAGGAGCTTGCTGTAAACAAAAATAATAAAAAATCACAATTTTTTTTT 1448
Db 3755 GAAGACCCCTGTGATGTCGTGTTGTTTAAACCCAGTTTA---CAGTAAATTTTTTTT 3810

QY 1449 TTTTTTTTGAAAAAGAGTCTTGCTCTTTTACCCTGGCTGGAGGCGAGTGGTGTGATCTCA 1508

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 14370.8 Seconds

(without alignments)

8124.892 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3910

Perfect score: 3910

Sequence: 1 aggaattccatttaaat.....cggacacggtcaccgctgc 3910

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_estum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hrc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hrc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620.4	15.9	1238	12	BM563325 AGENCOURT
2	491.4	12.6	579	12	BI559696 603252566
3	450.2	11.5	471	28	AQ760013 HS_3202_A
4	294.2	7.5	1910	11	BC035771 Homo sapi

5	290.6	7.4	2076	11	BC035510
6	290.6	7.4	2772	11	BC038630
7	286	7.3	754	12	BI827092
8	283.4	7.2	2772	11	BC038630
9	273.4	7.0	2076	11	BC035510
10	270	6.9	749	13	BU175345
11	269.4	6.9	1910	11	BC035771
12	259.4	6.6	929	9	AL580585
13	259.2	6.6	693	13	BU616025
14	255.8	6.5	2971	28	AF101960
15	255.2	6.5	958	9	AL578531
16	254.4	6.5	764	13	BU617735
17	252.4	6.5	940	13	BX424350
18	249.4	6.4	524	10	AW973992
19	249.2	6.4	580	12	BQ053726
20	249	6.4	978	13	BX325356
21	248.4	6.4	1201	13	BX385602
22	248.2	6.3	731	9	AI687343
23	247.2	6.3	775	12	BM905333
24	246.8	6.3	521	13	BU617227
25	246.8	6.3	598	12	BQ017808
26	246.8	6.3	616	13	BU617236
27	246.2	6.3	508	28	AQ418545
28	246.2	6.3	958	9	AL578531
29	246.2	6.3	969	13	BX415714
30	245.6	6.3	2284	11	BC036238
31	245.4	6.3	467	28	AQ388230
32	245.4	6.3	785	12	BM721013
33	244.8	6.3	663	29	AG174449
34	244.6	6.3	750	12	BM678642
35	244.2	6.2	1201	13	BX385602
36	244	6.2	548	14	CB144391
37	243.8	6.2	755	28	BZ599689
38	243.8	6.2	2971	28	AF101960
39	243	6.2	826	28	BZ611620
40	243	6.2	929	9	AL580585
41	242.6	6.2	493	10	AW328331
42	242.4	6.2	532	12	BG830539
43	242.4	6.2	861	12	BG676827
44	242.2	6.2	390	13	BU588888
45	242.2	6.2	759	13	BQ436286

ALIGNMENTS

RESULT 1
BM563325
LOCUS
DEFINITION AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.
ACCESSION BM563325
VERSION BM563325.1 GI:18810153
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1238)
NH-MGC http://mgc.ncbi.nlm.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12764 row: h column: 17
High quality sequence stop: 622.

RESULT 2
B1559696
LOCUS B1559696 579 bp mRNA linear EST 05-SEP-2001
DEFINITION G03252565F1 NIH MGC 97 Homo sapiens cDNA clone IMAGE:5294921 5'

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nhgri.nih.gov

Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
MCDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgenev,C., Vogt,J.B., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: P Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..2076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5094058"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH_MGC_8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 7.4%; Score 290.6; DB 11; Length 2076;
Best Local Similarity 67.2%; Pred. No. 2.9e-35;
Matches 498; Conservative 0; Mismatches 214; Indels 29; Gaps 5;
QY 1421 AAAATPATAAATAATACAAATTTTTTTTTTTTGGAGAAAGATCTTGCTCTTTTACC 1480
DB 741 ACAACGGTGAATAGATCTTTTTTTTTTTTGGAGATGGATTTTCACTCTGTAC 800
QY 1481 CTGGCTGAGGCGAGTGTGTGATCTAGCTACGTACGCAACTTGCCTCCCGGGTTCAAG 1540
DB 801 CAGGCTGGAGTGCAATGGCGTGATCTTGGCTCACCAGAACCTCCGCTCTCTGGGTTCAA 860
QY 1541 CAATTTCTCTGCTCAGCTCCAGCTCCAGTACGTGGGACATCAGGCACTTCCACCATGCCA 1600
DB 861 CAATTTCTCTGCTCAGCTCCAGCTCCAGTACGTGGGATTAAGGCAATGACCCCAAAACCG 920
QY 1601 GCTGATTTTGTATTTTGTAGATGGGATTTTCACTTTGTGGCCAAAGTGTCTCAA 1660
DB 921 GAT-AAATTTTGTATTTTGTAGATGGGATTTTCACTTTGTGGCCAAAGTGTCTCAA 979
QY 1661 ACTTTTGTCTCATATTTGTGTAATTTGTTTGTCTTTTGTCTGAGGTAGGCGCCCGCA 1720
DB 980 ACT-----CCGACCTCTGCTGATCCGATGCTGCTGCGCTCCCAAGTGTGGGATT 1031
QY 1721 CCAAAAAAATAATCTTAGAATCAAAATCAGTGTGTGTGTGTGACCACTGTCACTTTGAG 1780
DB 1032 ACATGCGGTGACCCACCGTCCCGACATAGATGCTTTTAATTTTCTGGAGGAAAAAGCA 1091
QY 1781 AACCACTGTGACACAGCGCTCAGGAGTAGAGTGATCTCTGTCGAAAGAGAAATAGA 1840
DB 1092 AAGCAAAAGAACAGCTGGATATTTTAAAG-----ACTAAAGAGGAAAAACAA 1137
QY 1841 ATGAAATATTTCTCGGCGCAGGCGTGTGTGCTC-ATGCTGTATATCCACCATTTTGGG 1899
DB 1138 AAAGGAGATAGACAGCGCCAGACGTGTGTGCTCAACGCTGTGTATCCGACATTTTGGG 1197
QY 1900 AGGCGAAGGCGATGGGATCACCTGAGGTACAGAGTTTCAAAACAGCGCTGGGCCAACATGGT 1959
DB 1198 AGGCGGAGGAGGTGGATCACTGAGGTACAGAGTTTCAAGACAGCGCTGACCAACATGGT 1257
QY 1960 GAACCCCGCTCTTACTAAATAACAAAAATAGCTAAGTGTGGTGGCGCATCCCTGTA 2019
DB 1258 GAAACCCCTGTTTC-----AAAATACAAAAAATTAGCTGGGCGTGGTGGCGGCACCTGTG 1312

QY 2020 ATCCAGCTACTTGGAGGCTGAGCGAGAGAAATTTCTTGAACCCGGAGGAGAGGTTG 2079
DB 1313 ATCCAGCTACTTGGAGGCTGAGCGAGAGAAATCCCTTGAACCCAGGAGGTGGAGTTG 1372
QY 2080 CAGTGAAGGAGATCACACCTGCTCAGCTCCAGCTGGGGAGAGAGAGACTTCTCTC 2139
DB 1373 CAGTGAAGGAGATCACCGCATTTGACTCCAGCTGGGCGACAGTAAAAAACTCCATCT 1432
QY 2140 AAAAAACAAAAACAAAAAGA 2160
DB 1433 CAAAAAAGGAGATA 1453

RESULT 6

BC038630/c

LOCUS BC038630 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, Similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.

ACCESSION BC038630

VERSION BC038630.1 GI:24116283

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2772)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL

Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadan@systemsbio.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 73 Row: 1 Column: 4

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers

1..2772

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:5263792"

/tissue_type="Brain, hippocampus"

/clone_lib="NIH_MGC_95"

/lab_host="DH10B"

/note="Vector: pBluescript"

ORIGIN

Query Match 7.4%; Score 290.6; DB 11; Length 2772;

Best Local Similarity 69.5%; Pred. No. 2.4e-35;

Matches 508; Conservative 0; Mismatches 189; Indels 34; Gaps 7;

QY 1441 TTTTTTTTTTTTTTTGGAGAAAGTCTTCTTTTCACCTGGCTGGGCGGAGTGGT 1499

DB 2753 TTTTTTTTTTTTTTTGGTAAAGACGGAGTTTCTTTCTCCAGGCTGGAGTGGT 2694

QY 1500 GTGATCTCAGTCACCTCAACTTTCGGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGCC 1559

DB 2693 GCCATCTCGGCTCGCTGCAAACTGCACCTCCGGGTTCAACAGATTTCTCTGCTTCAGCC 2634

```

QY 1560 TCCCAAGTAGTGGGACTACAGGCACTTCCACATGCCAGCTGATTTTGTATTTT 1619
Db 2633 TCCCAAGTAGCGGACACAGGTGGCGGCACACGCCCGCAATTTCTGTATATTTA 2574
QY 1620 GTAGAGATGGGATTTCACTTTTGTGGCCAAAGCTGGTCTCAAACTTTTGTGTCAATAAT 1679
Db 2573 GTATATTGGGGTTTACCATGTGTGGCCAGGCTGGTCTCAAACTCTCTGAGCTCAGGTGAT 2514
QY 1680 GTTGTAACTATTGTTCTTTTGTCTGAGGTAGGCGCCCGACACCAAAAAAATAAATTTA 1739
Db 2513 CCATCTGCCCTCAGCCTCCCAAGGTGGTG-----GGAATACAGGCATGAGCCACA 2465
QY 1740 GAATCCAAATCAGTGTGTGGTTTGACCACTGTCTGAGAACCCAGAGTGTGACAGGG 1799
Db 2464 GCTCCAGCTAGGATTTCTGATATATCACTATCTGCTGTGTTAGA-AGTGAACCTGGC 2406
QY 1800 CCTCAGAGTAGAGTGTATCTCTGCTCGAAGAGAAATAGAATGAAATATTCTCCGGGC 1859
Db 2405 TAAATTAATAGAGATGAATTTACTGAAAAGTGTATA-----TGC 2364
QY 1860 CAGCGTGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCAAGGCATGTGATCA 1919
Db 2363 CAGAGTGGTGGCTCATGCTGTAAATCACAACACCTTGAGAGCCGAGGCGGTAGGTCA 2304
QY 1920 CCTGAGTCAAGGAGTTCAAAACCAAGCTGTGCCCAACATGTTGAAACCCCGTCTCTACTAAA 1979
Db 2303 CCTGAGTCAAGGAGTTC-AAACCAAGCTGTGACCAATGTTGAAACCTTGGCTCTACTAAA 2245
QY 1980 AATCAAAAAATAGTAAAGTGTGGTGGCCATGCTGTAAATCCAGCTACTTTGGAGGG 2039
Db 2244 AATAC-AAAAATCAGCAGGCTGAGTGGTGCATGCTGTAGACCAGCTACTCCGAGGC 2186
QY 2040 TGAGGCAGAGATTTCTTGAACCCGAGGAGGAGAGTTCAGTGAAGCAGATCACACC 2099
Db 2185 TGAGGCAGGAGAAATGCTTGAACCCAGGAGGTAGAGGTTCAGTGAAGCAGATCACACC 2126
QY 2100 ACTGCATCTCAGCTCGGGGAGAGAGAGTTCCTCTCAAAAAACAAAAACAAAG 2159
Db 2125 ATTGCACT-CAGCTCGGCAACAGAGAGAGTTCATCTCAAAAAAAGAAAAAGAA 2067
QY 2160 AATTAAGCAA 2170
Db 2066 AAAAAAGAAA 2056

```

```

RESULT 7
BI827092
LOCUS 603075763F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167695 5',
DEFINITION mRNA sequence.
ACCESSION BI827092
VERSION BI827092.1 GI:15938629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 754)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloned distribution: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW1416 row: n column: 16
High quality sequence stop: 741.
Location/Qualifiers

```

```

source 1. .754
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5167695"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
Note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

```

ORIGIN

```

Query Match 7.3%; Score 286; DB 12; Length 754;
Best Local Similarity 99.7%; Pred. No. 2.7e-34;
Matches 297; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 3613 GGCAAGCAGAGTGGCAACCAATTCCTTCGGCCAATGACGAGCCGAGTTTACAGAAGC 3672
Db 1 GGCAAGCAGAGTGGCAACCAATTCCTTCGGCCAATGACGAGCCGAGTTTACAGAAGC 60
QY 3673 CTCATTAGCATTTCCCAAGCAGGCGAGGGGAGAGCCGGGGTGGTGTGTGTCGGT 3732
Db 61 CTCATTAGCATTTCCCAAGGCA-GGGCAAGGCGAGAGCCGGGGTGGTGTGTGTCGGT 119
QY 3733 GTCGCGAGCATCCCGGGCCCTGCTCGGTCGCGGAGCCTCGGCTCTGTCCTCC 3792
Db 120 GTCGCGAGCATCCCGGGCCCTGCTGCGGTGCGCGAGCCTCGGCTCTGTCCTCC 179
QY 3793 CCTCCCGCCCTTACCTCCACGCGGAGCCGCGCCAGTCAACCTCGCACTTTGCC 3852
Db 180 CCTCCCGCCCTTACCTCCACGCGGAGCCGCGCCAGTCAACCTCGCACTTTGCC 239
QY 3853 CTGCTGTGCGAGCGGATAAAGGGGGCTGAGAAATACCGGACAGGTCAACCGTTGC 3910
Db 240 CTGCTGTGCGAGCGGATAAAGGGGGCTGAGAAATACCGGACAGGTCAACCGTTGC 297

```

RESULT 8

```

BC038630
LOCUS 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.
ACCESSION BC038630
VERSION BC038630.1 GI:24116283
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2772)
JOURNAL Strausberg, R.
COMMENT Direct Submission
Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org

```

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 73 Row: i Column: 4
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..2772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5263792"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 7.2%; Score 283.4; DB 11; Length 2772;
Best Local Similarity 66.1%; Pred. No. 3.1e-34;
Matches 486; Conservative 0; Mismatches 216; Indels 33; Gaps 4;
QY 1441 TTTTCTTTTCTTTTGTGAGAGAGCTCTGCTCTTCCACCTGGCTGAGGGCAGTGGTG 1500
Db 2071 TTTCTTTTCTTTTGTGAGATGGAGTCTTCTCTGTTGCCAGGCT-GAGTGCATGGTG 2129
QY 1501 TGATCTAGCTCACTGCAACTTTCGCTCCCGGTTCAAGCAATCTCTGCTTCAGGCT 1560
Db 2130 TGATCTCGGCTCACTGCACTCTTACCTCTGCTGGTTCAAGCAATCTCTGCTTCAGGCT 2189
QY 1561 CCCAAGTAGCTGGGACTACAGCACTTCCCAACCAAGCCCAAGTATTTGTATTTTATG 1620
Db 2190 CCGAGTAGCTGGGTCTACAGCATGACCACTAGCCCTGGCTGATTTTGTATTTTATG 2249
QY 1621 TAGAGATGGGATTTACATTTGTCGCAAGCTGGTCTCAAACTTTTCTGTCATATG 1680
Db 2250 TAGAGGCAAGGTTTACCATGTTGCTCAGGCTGTTGAACCTCTCAGCTCAGGTGACCT 2309
QY 1681 TTGTAACATATGTTCTTCTGCTGAGGTAGGCCCCCAGACCAAAAAATAATCTTAG 1740
Db 2310 ACCCGCTCGGCTCTCAAGTGTGTGATTACAGGCAATGACCACTCTCTGGCATAT 2369
QY 1741 AATCCAAATCAGTGTGTGTTGACCACTGTCACTTGAGAACCAACAGTGTGACCAAGGC 1800
Db 2370 AACACTTTTCAGTAAAT-----CATCTTCTATTTAATTTAGCCAGATTCACCTTA 2422
QY 1801 CTCAGGAGTAGAGTGATCTCTGCTCGAAGAGAAATAGATGAATATTTCTCGGGCC 1860
Db 2423 ACAAGCAGATAGTTGATATATCAAGAA-----TCTTAGCT 2458
QY 1861 AGCGTGTGCTCATGCTCTAATCCAGACATTTGGAGGCCAAGCATGTGGATCAC 1920
Db 2459 GGGAGCTGTGCTCATGCTGTATATCCACACCTTGGAGGCTGAGCGAGTGGATCAC 2518
QY 1921 CTGAGGTCAGAGTTCAAAACAGCTGTGGCAACATGTTGAAACCCCGTCTCTACTAAAA 1980
Db 2519 CTGAGGTCAGAGTTTGAGACCAAGCTGGCCAAACATGTTGAAACCCCAATATACTAAAT 2578
QY 1981 ATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTATATCCAGCTACTTGGAGGGT 2040
Db 2579 ATAC-AGAAATTTGCTGGGCGTGTGGCGGCCACCTGTGTGCTCCGGTACTTGGAGGCT 2637
QY 2041 GAGCAGGAGAAATTTCTTGAACCCCGGAGGAGAGGTTGAGTGAAGCGAGATCACACCA 2100
Db 2638 GAGCAGGAGAAATCTGTTGAACCCCGGAGGTTGAGTGAAGCGAGATGGACCA 2697
QY 2101 CTGCACTCCAGCTGGGGGAGAGCGAGATCTCTCTCAAAAAACAAAAACAAAAA 2160
Db 2698 TTGCACTCCAGCTGGGAGAGCAAGAGCAAAACTCCGCTTACCAAAAAACAAAAAAGA 2757
QY 2161 ATTAAGCAAAATAGA 2175
| | | | |

Db

2758 AAAAAAAAAAAAAA 2772

RESULT 9

BC035510/c

LOCUS

BC035510

DEFINITION

Homo sapiens sialophorin (gp115, leukosialin, CD43), mRNA (cdna

clone IMAGE:5094058), with apparent retained intron.

BC035510

BC035510.1

GI:23271054

HTC

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2076)

Strausberg, R.L., Feingold, Z.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheerz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2076)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-x@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>Contact: nisc.mgc@nih.gov

Akhtar, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,

Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 39 Row: p Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.


```

Db      450 GCGTGGTGGTCTATGCGCTGTAAATCTGGCACTTTGGGAGGCCAAGGCGAGCATCACTG 509
QY      1923 GAGGTGAGGAGTTCAAAACCCAGCGCTGCCAACATGTGAACCCCGCTCTCTCTAAAT 1982
Db      510 GAGCCCGAGGATCAAGACCGCGCTGGGCAATAGCAAAACCTTGTCTC--CAAAAAAT 567
QY      1983 AAAAAAATAGTAAGTGTGGTGGCGCATGCGCTGTAATCCCAAGTACTTTGGGAGGGTGA 2042
Db      568 ACAAAAAAATAGCTAGATGGTGTGTGTGCACTGCTAGCTACTCAGGAGTCTGA 627
QY      2043 GCGCAGGAGATTTCTTGACCCCGGAGGCGAGAGTTCAGTGAAGGAGATCACACACT 2102
Db      628 GTTGGGAGGATGCTTTGAGTCAGGAGTGTGAGGGTATAATGAACCTATGGGTGCACAACT 687
QY      2103 GCACCTCAGCGTGGGCGAGAGCGAGACTTCTCTCAAAAAAACAACAAAAA 2158
Db      688 GTACTCTAGTGGGACACAGAGGAGAGCCCTGCTCCCAAAAAAANAANAANA 743

RESULT 11
BC035771
LOCUS   1910 bp mRNA linear HTC 19-NOV-2003
DEFINITION Homo sapiens RAD1 homolog (S. Pombe), mRNA (cDNA clone IMAGE:5582524), with apparent retained intron.
ACCESSION BC035771
VERSION   1
KEYWORDS  BC035771.1 GI:23958889
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1910)
AUTHORS   Klausner,R.D., Collins,F.S., Wagner,L.H., Derge,J.G.,
Altschul,S.F., Zeeberg,B., Butow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Ustin,T.B., Toehiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,T.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,A., Schein,J.E., Jones,S.J., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 1910)
Strausberg,R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov

```

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Lari, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 79 Row: d Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: retained intron.

Location/Qualifiers

source

```

1. .1910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5582524"
/tissue_type="Testis, embryonal carcinoma"
/clone_lib="NTH MGC_92"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

ORIGIN

```

Query Match      6.9%; Score 269.4; DB 11; Length 1910;
Best Local Similarity 64.8%; Pred. No. 5,4e-32;
Matches 471; Conservative 0; Mismatches 236; Indels 20; Gaps 4;

QY      1441 TTTTATTTTATTTTGGAGAGAGTCTTCTCTTTCACCCCTGGCTGGAGGCGAGTGTG 1500
Db      451 TTTATTTTATTTTGGAGAGAGATCTTCTCTTTCACCCCTGGCTGGAGTACATGTGA 510
QY      1501 TGAATCTAGCTACCTGCAACTTTCGCTCCCGGGTTCGAAGCAATCTCTCTGCTCAGCT 1560
Db      511 GGATTTACAGTTTCACTCCAGCGCTCAATCTCTAGGCTCAAGCAATCTCTCCATCTCAGCTT 570
QY      1561 CCCAAGTAGCTGGAGTACAGGCACTTCCCAACCATGCCAGCTGATTTTGTGA-TTTTTA 1619
Db      571 CCCAAGTAGCTACAGTACAGGTCATGCTCCATGCCCTGCTAAATTTTGTGATTTTTT 630
QY      1620 GTAGAGATGGAGTTTCACTTTTGGTGGCAAGCTGGTCTCAAACTTTTGTGTGCTAAT 1679
Db      631 GTAGAGATGGGGTTTACTATGTTCCCTAGGCTAATCTTGAAC---TCTTGAGCTCAAG 686
QY      1680 GTTGTAACTATTGTTCTTCTTGTGAGTAGGCGGCCCGCCAGCAAAATAAATCTTA 1739
Db      687 CAATCCACCGCTCGGCTCCCAAGGGCTGGGATTTACAGGGCTGAGCTATCATGCTG 746
QY      1740 GAATCCAAATCAGTGTGTGTGGTTTGAACCTGTCTGAGAACCCAGCTGTGACCAAGG 1799
Db      747 GCCTAATGTTATTTTAAAGTTGAATCTAAAGAAATTTAAACAAATATAATCAGCT 806
QY      1800 CCTCAGAGTAGAGGTGATCTCTGCTCGAAGAGAGAAATAGATAAATAATCTCCGGG 1859
Db      807 CAGTAAGGATGGATATTTTGTACATTTTAAACCTA-----TTATTTTAGGC 852
QY      1860 CAGCGTGGTGGCTCATGCTCTTAATCCCAAGCTTTGGGAGCCAGGCGATGGATCA 1919
Db      853 CGGAGTGGTGGCTCACACCTGTAATCCAGCATTTTGGGAAGCCGAGGCGGGGATCA 912
QY      1920 CCTGAGTCAAGG-AGTTCAAAACAGCTGGCCCAACATGTTGAAACCCCGCTCTCTACTAA 1978
Db      913 TCTGAGTCAAGGAGTTGAGATCAGCTTGGCCCAACATGTTGAAACCCCTCTACTA 972
QY      1979 AATACAAAAAATAGTAAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTGGAGG 2038
Db      973 AAAATCAAAAAATAGCTGGGTGTGGTGGTGTCTGTAATCCAGCTACTTGGAGG 1032
QY      2039 GTGAGGCGAGGAGATTTCTTTGACCCGGGAGGCGAGGTTGTCAGTGAAGCGAGATCACAC 2098

```



```

||||| 1033 CTGAGGCAAGAGATCACTTGAACCTGGAAGCGGGGTTCAGTGAGCCGAGATTGCAC 1092
|||||
QY 2099 CACTGCACTCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAA 2158
|||||
Db 1093 CATTGCACTCCAGCTGGGTACAGAGCAAGACTCCGCTCTCAAAAAACAAAAAATTAAA 1152
|||||
QY 2159 GAATTA 2165
|||||
Db 1153 ATAAAA 1159
|||||

RESULT 12
AL580585 929 bp mRNA linear EST 01-JUN-2003
LOCUS AL580585 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ015YK20 3-PRIME, mRNA sequence.
ACCESSION AL580585
VERSION AL580585.2 GI:31318853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12946748.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8901.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015BF10NP1&cluster=8901.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015BF10NP1.
Location/Qualifiers
1. .929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YK20"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 6.6%; Score 259.4; DB 9; Length 929;
Best Local Similarity 62.2%; Pred. No. 2.9e-30;
Matches 507; Conservative 14; Mismatches 245; Indels 49; Gaps 6;

QY 1441 TTTTITTTTTTTTTTGAAAGAGTCTGCTTTCACCGCTGGCTGGAGGCGAGTGGTG 1500
|||||
Db 116 TTTTITTTTTTTTTTGATCGAKTCTCTCCCTGTGCGGAGGTCAGACAGTGGCG 175
|||||

QY 1501 TGATCTCAGCTCACTGCAACTTTCGCTCCCGGTTTCAAGCAATTCCTGCTCAGCT 1560
|||||
Db 176 TGATVTCAGCTCGTGCACTCTGCTCCCGTGTTCAGCAATTTCTGCTCAGCT 235
|||||
QY 1561 CCCAGTAGCTGGGACTACAGGCACTTCCCAATGCCAGCTGATTTTGTATTTTAT 1620
|||||
Db 236 CCCGAGTWTGCTGGGACTACAGGCGGTGTCACACACTCAGCTAATTTTGTATTTCCAG 295
|||||

```

```

QY 1621 TAGAGATGGGATTTTCACTTTTGTGGCCAAAGCTGCTCTCAACATTTTTCCTGCTCATATTG 1680
|||||
Db 296 TAGAGACTG--TTTACATGTTTGTGTCAGGCTGCTGTGAACCTCTGACTCTGTTGA---- 349
|||||
QY 1681 TTGTAATAATTTGTTTCTGAGGTAGGGCCCCCAGACCAAAAAAATAAATCTTAG 1740
|||||
Db 350 -----TCGCCCCACCTCAGCCTCCCAAGTGTGGATTTACAGGCGAG 392
|||||
QY 1741 AATCCAAATCAGTGTGTTGTTGACCACTGTCTCTCGAAAGAGAAATAGAAATATTTCTCGGGGCC 1800
|||||
Db 393 CCACC-----TTGCCCGGCCGAAGTTAAAGATGCAGTTTGGGATAGAC 437
|||||
QY 1801 CTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCGGGGCC 1860
|||||
Db 438 ATCATAGGATATTTCTTCCCTT-----TACTCATAAAGTTGTCTGAGCTTTGGCC 489
|||||
QY 1861 AGCGGTGGTGGCTATGCTGCTGTAATCCAGCACTTTGGGAGGCCAAGCATGTGATCAC 1920
|||||
Db 490 GGGCGCGATGGCTCAGCCTGTTATCTAGTACTTTGGGAGGCTGAGCGGCGAGATCAC 549
|||||
QY 1921 CTGAGGTCAGGAGTTCAAAACACAGCCTGGCCCAACATGTTGAAAACCCGCTCTTACTTAAAA 1980
|||||
Db 550 CTGAGGTCAGGAGTTTGTGACCAAGCCTGGCCCAACATGACAAAACCCCTGTCTTACTTAAAA 609
|||||
QY 1981 ATACAAAAAATTAGCTAAGTGTGCTGGCGCATCGCTGTATATCCAGCTACTTGGGAGGGT 2040
|||||
Db 610 ATAC--AAAACTAGCCAGGTGTGTTGGTGGCACCTGCAGTCCCACTATTTCGGAGGCT 667
|||||
QY 2041 GAGGCGAGGAGAAATTTCTTGAACCCGGGAGGAGAGGTTGCAAGGAGGAGATCACACCA 2100
|||||
Db 668 GAGGCGAGGAGAAATCACCCGAAACCTGGGAGGAGAGGTTGTAGCGAGCCGAGATCTTGCCA 727
|||||
QY 2101 CTGCACTCCAGCTGGGCGAGAGAGCGAGCACTTCTCTCAAAAAACAAAAACAAAAAGA 2160
|||||
Db 728 TCACATCCAGCTGGAGGACAGAGCAAACTGTGTCACAAAAA--AAAAAAGT 786
|||||
QY 2161 ATTAAGCAAAATTAGCATTCAGAGAGAACCTGAAGGGGCTCAGACCACTACAGATTTTC 2220
|||||
Db 787 TGTAGCTTAAATAATCCCTTCACGAGAGCTTCTTTGCTCTTAAAMCTTTCTCTTTT 846
|||||
QY 2221 TGTGCCACATGCCAAGTACTTCTGAGGCGATGACTG 2255
|||||
Db 847 TTTAAWCTGTYTGCGKGKTTTGAGGGGAAACAG 881
|||||

RESULT 13
BU616025 693 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-DF0-ben-m-20-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone
DEFINITION UI-H-DF0-ben-m-20-0-UI 3', mRNA sequence.
ACCESSION BU616025
VERSION BU616025.1 GI:23282240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 693)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 9-305, >ALU (matched complement) 326-387, >L1MB7#LINE/L1
(matched complement) 388-676, >ALU

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 24113.2 Seconds
(without alignments)
7114.428 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattacatttaaat.....gggaactccacgcacgcggg 3958

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3958	100.0	4150	6	AX528729	AX528729 Sequence
2	3958	100.0	4150	9	AF320307	AF320307 Homo sapi
3	3928.2	99.2	159351	2	AC018783	AC018783 Homo sapi
4	3917.8	99.0	115278	9	AL139819	AL139819 Human DNA
5	1827.2	46.2	185644	2	AL391723	AL391723 Homo sapi
6	1799.8	45.5	1970	6	AX431388	AX431388 Sequence
7	1758	44.4	161827	2	AC139668	AC139668 Papio ham
8	1448.2	36.6	179879	2	AC026883	AC026883 Homo sapi
9	816.6	20.6	190784	2	AC139669	AC139669 Lemur cat
10	521	13.2	17088	4	AY241932	AY241932 Bos tauru
11	498.4	12.6	617	6	AX357255	AX357255 Sequence
12	460	11.6	164201	9	AL157935	AL157935 Human DNA
13	447.2	11.3	200288	2	AC119725	AC119725 Homo sapi
14	433	10.9	175842	9	AL731547	AL731547 Human DNA
15	423.8	10.7	190508	9	AC005037	AC005037 Homo sapi
16	415.2	10.5	44496	9	AC004760	AC004760 Homo sapi
17	415.2	10.5	190814	9	AC006111	AC006111 Homo sapi
18	408.8	10.3	183690	9	AC091180	AC091180 Homo sapi
19	407.4	10.3	170528	9	AC083866	AC083866 Homo sapi
20	407.4	10.3	217615	9	AC011005	AC011005 Homo sapi
21	404.2	10.2	137289	9	AL161909	AL161909 Human DNA
22	403	10.2	189179	2	AC090265	AC090265 Homo sapi
23	402.8	10.2	181597	9	AC090515	AC090515 Homo sapi
24	402.8	10.2	187734	9	AC025918	AC025918 Homo sapi
25	401.8	10.2	204963	2	AC145617	AC145617 Homo sapi
26	394.2	10.0	179599	2	AC13278	AC13278 Homo sapi
27	394.2	10.0	195558	9	AC007448	AC007448 Homo sapi
28	392.2	9.9	155559	2	AL133336	AL133336 Homo sapi
29	392.2	9.9	169835	9	AP005717	AP005717 Homo sapi
30	391.6	9.9	171641	9	AC018512	AC018512 Homo sapi
31	390	9.9	139632	9	AC145334	AC145334 Pan trogl
32	390	9.9	171555	2	AC087537	AC087537 Homo sapi
33	390	9.9	181891	9	AC023356	AC023356 Homo sapi
34	389.8	9.8	67199	9	AC025162	AC025162 Homo sapi
35	388.4	9.8	172421	2	AC013279	AC013279 Homo sapi
36	386.2	9.8	112364	9	AC109927	AC109927 Homo sapi
37	386	9.8	73162	9	AC105095	AC105095 Homo sapi
38	386	9.8	197092	2	AC036155	AC036155 Homo sapi
39	384.8	9.7	140529	2	AC011486	AC011486 Homo sapi
40	384	9.7	133475	9	AC006254	AC006254 Homo sapi
41	382.4	9.7	189639	9	AC010997	AC010997 Homo sapi
42	381.6	9.6	80775	9	AC068539	AC068539 Homo sapi
43	381.2	9.6	150681	9	AC006011	AC006011 Homo sapi
44	381.2	9.6	187064	9	AC011476	AC011476 Homo sapi
45	381.2	9.6	188413	2	AC145834	AC145834 Homo sapi

ALIGNMENTS

RESULT 1
AX528729
LOCUS AX528729
DEFINITION Sequence 15 from Patent WO0236780.
ACCESSION AX528729
VERSION AX528729.1 GI:25172800
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Prouty, S.M., Zhang, L. and Stenn, K.S.
TITLE Stearoyl-coa desaturase gene promoter
JOURNAL Patent: WO 0236780-A 15 10-MAY-2002;

Db	1981	ATACAAAAATTAGCTAAGTGTGGTGGGCGATGCTGTAATCCAGAGCTTACTTGGGAGGGT	2041
Qy	2041	GAGGCAGGAGAATTTCTTGAACCCGGGAGGACAGAGTTGCAGTGNAGCGAGATCACACCA	2101
Db	2041	GAGGCAGGAGAATTTCTTGAACCCGGGAGGACAGAGTTGCAGTGNAGCGAGATCACACCA	2100
Qy	2101	CTGCATCTCAGCCTCGGGGAGAGCGAGACTTCTCTCAAAAAACAACCAAGG	2160
Db	2101	CTGCATCTCAGCCTCGGGGAGAGCGAGACTTCTCTCTCAAAAAACAACCAAGG	2160
Qy	2161	ATTAAAGCAATTAGACATTCGAGAGAGAACCTGTAAGGGGTGCAGCACGTAAGATTTC	2220
Db	2161	ATTAAAGCAATTAGACATTCGAGAGAGAACCTGTAAGGGGTGCAGCACGTAAGATTTC	2220
Qy	2221	TGTGCCCATGTCCTGAGGATGACTTCTGAGGCATGACTGAGTGAAGTGTGCACATCTGAATCA	2280
Db	2221	TGTGCCCATGTCCTGAGGATGACTTCTGAGGCATGACTGAGTGAAGTGTGCACATCTGAATCA	2280
Qy	2281	TCCAGTCTTGTTCAGAACTTTTTCACACGGGACAGGGAGCCAGGACTGGAATGCAATCTCCT	2340
Db	2281	TCCAGTCTTGTTCAGAACTTTTTCACACGGGACAGGGAGCCAGGACTGGAATGCAATCTCCT	2340
Qy	2341	GGTCACTGCGCCAGAGAGTTGGCTTGACCTTGACACCCAGTGGCCACAAGAGAGCTGCTT	2400
Db	2341	GGTCACTGCGCCAGAGAGTTGGCTTGACCTTGACACCCAGTGGCCACAAGAGAGCTGCTT	2400
Qy	2401	AGTCTACCTCCACAGGAATCCAGGTGCTGTCTTCTCGGAAATGAATCATTTGGGCGAG	2460
Db	2401	AGTCTACCTCCACAGGAATCCAGGTGCTGTCTTCTCGGAAATGAATCATTTGGGCGAG	2460
Qy	2461	CACCTCGTATTTTCTCTTCTCCAGGGGAGGATCCTAGGGCAGTATTTGGGAAAGACA	2520
Db	2461	CACCTCGTATTTTCTCTTCTCCAGGGGAGGATCCTAGGGCAGTATTTGGGAAAGACA	2520
Qy	2521	TGGGCATGGAAGGACACCGGTGGAATGCATAGACCTGCCTGGTTCTGAGCTCTCATGTTAA	2580
Db	2521	TGGGCATGGAAGGACACCGGTGGAATGCATAGACCTGCCTGGTTCTGAGCTCTCATGTTAA	2580
Qy	2581	GGCTCTTCACAGACACGGAAGAATGGGGGCACAGGGACAGATCAGTATGGGTTCAGAGCATC	2640
Db	2581	GGCTCTTCACAGACACGGAAGAATGGGGGCACAGGGACAGATCAGTATGGGTTCAGAGCATC	2640
Qy	2641	TCAGGACCGGCGCAATATGGTCTTGACGAGGGATTAAGAGCTTGGGTCTCATATGGT	2700
Db	2641	TCAGGACCGGCGCAATATGGTCTTGACGAGGGATTAAGAGCTTGGGTCTCATATGGT	2700
Qy	2701	GTCTTGGGCTCAACTGCCAGCTCCGTCACTTACTGTTGCTGACCAATGGGCAAGTTA	2760
Db	2701	GTCTTGGGCTCAACTGCCAGCTCCGTCACTTACTGTTGCTGACCAATGGGCAAGTTA	2760
Qy	2761	TTCCATCTCTCCATATCTCTTCTTCTTCAATGGAATAATGGGGTACCCACCTCC	2820
Db	2761	TTCCATCTCTCCATATCTCTTCTTCTTCAATGGAATAATGGGGTACCCACCTCC	2820
Qy	2821	CAGGTCACAGAGAGCTTACAGAAAAAGATTCTTGTAATGGCTTGGCAGTAATAATTC	2880
Db	2821	CAGGTCACAGAGAGCTTACAGAAAAAGATTCTTGTAATGGCTTGGCAGTAATAATTC	2880
Qy	2881	AATACCTGCCAGCTATCTTATTTCACATCCAAAGCCCTTGCCTGCTGCTGGGTGAAAA	2940
Db	2881	AATACCTGCCAGCTATCTTATTTCACATCCAAAGCCCTTGCCTGCTGCTGGGTGAAAA	2940
Qy	2941	CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAATTTACAACCTGCCAG	3000
Db	2941	CACATGTCAGTGTTCCTGACGGTTTCCACAAAGAGATTTCCAAATTTACAACCTGCCAG	3000
Qy	3001	TCTTGAGAAATCTCCAAAAATCCCGCAGCATCTCTGGAGGGCGGGCTTGGGGATGGGAC	3060
Db	3001	TCTTGAGAAATCTCCAAAAATCCCGCAGCATCTCTGGAGGGCGGGCTTGGGGATGGGAC	3060
Qy	3061	TGCCCCCGGGTCTCTGACAGATGCGTGGCGGAGGACACACACACAGCAGCGCTG	3120
Db	3061	TGCCCCCGGGTCTCTGACAGATGCGTGGCGGAGGACACACACACAGCAGCGCTG	3120

QY	3121	TGTGTGCGCGCGAGTCTCCGGTGC	CGTCTCCCGGTGAGCAGCGGTGGTGGCGCGGGC	3180
DB	3121	TGTGTGCGCGCGAGTCTCCGGTGC	CGTCTCCCGGTGAGCAGCGGTGGTGGCGGGC	3180
QY	3181	AGAGCCATTGTTCGACAGCGTACG	AGCCCCCGCGCTCGCCCCCGGAGGCGGGGCT	3240
DB	3181	AGAGCCATTGTTCGACAGCGTACG	AGCCCCCGCGCTCGCCCCCGGAGGCGGGGCT	3240
QY	3241	TCCCGCGTCCCCAAGCTCCAGATC	CTTGGGTGGTGCACGTCTCCCTGCACGCGCTG	3300
DB	3241	TCCCGCGTCCCCAAGCTCCAGATC	CTTGGGTGGTGCACGTCTCCCTGCACGCGCTG	3300
QY	3301	GGGGACCGGAGAGCGGACGAGATG	TATGTGTGGCGCCCCCGAGGGTTTCA	3360
DB	3301	GGGGACCGGAGAGCGGACGAGATG	TATGTGTGGCGCCCCCGAGGGTTTCA	3360
QY	3361	GTTTCTTGAGAAAATTCTCCCACT	CGCCACCGCTTCTCCGTGTCCCGAGGGCGGT	3420
DB	3361	GTTTCTTGAGAAAATTCTCCCACT	CGCCACCGCTTCTCCGTGTCCCGAGGGCGGT	3420
QY	3421	CCTGGGCTAGGCTCCGCGCCCGC	CAACCGGGTCCCAGCGCTTCCAGAGAGAAA	3480
DB	3421	CCTGGGCTAGGCTCCGCGCCCGC	CAACCGGGTCCCAGCGCTTCCAGAGAGAAA	3480
QY	3481	GCTCCCGACCGGGATCCGGGCAG	AGCCCGCGGGTGGAGAGAGCTGAGAGG	3540
DB	3481	GCTCCCGACCGGGATCCGGGCAG	AGCCCGGGTGGAGAGAGCTGAGAGG	3540
QY	3541	AGAAACAGAGGGAGGGGACGAG	AGCTGGCGGAGGGAGACAGATTTGCCCA	3600
DB	3541	AGAAACAGAGGGAGGGGACGAG	AGCTGGCGGAGGGAGACAGATTTGCCCA	3600
QY	3601	GCCAAATGGCAACGGCAGGACG	AGGTGGCAACCAATTTCCCTTCGGCCAT	3660
DB	3601	GCCAAATGGCAACGGCAGGACG	AGGTGGCAACCAATTTCCCTTCGGCCAT	3660
QY	3661	GTTTACAGAAAGCTCATTTAGCAT	TTTCCAGAGGCGGCGAGGCGAGGCGCGGTGG	3720
DB	3661	GTTTACAGAAAGCTCATTTAGCAT	TTTCCAGAGGCGGCGAGGCGAGGCGCGGTGG	3720
QY	3721	TGTGTGTGTGCGGACGATPCCCG	CGCCCTTCTCGGTTCGCGGAGCTTCGGCC	3780
DB	3721	TGTGTGTGTGCGGACGATPCCCG	CGCCCTTCTCGGTTCGCGGAGCTTCGGCC	3780
QY	3781	TCGTGCTCTCCCGCTCCCGCTT	ACCTTCAGCGGGACCGCCCGCGCAGTCA	3840
DB	3781	TCGTGCTCTCCCGCTCCCGCTT	ACCTTCAGCGGGACCGCCCGCGCAGTCA	3840
QY	3841	TGCACTTTTGCCCTGCTTTGGC	AGCGGATAAAGGGGGCTGAGGAAATAC	3900
DB	3841	TGCACTTTTGCCCTGCTTTGGC	AGCGGATAAAGGGGGCTGAGGAAATAC	3900
QY	3901	CACCCGTGCGAGCTTAGCTTTAA	ATTTCCGGTTCGGGACCTTCCACGACCGCGG	3958
DB	3901	CACCCGTGCGAGCTTAGCTTTAA	ATTTCCGGTTCGGGACCTTCCACGACCGCGG	3958

gene promoter: requirement of a conserved CCAAT cis-element
Biochem. J. 357 (Pt 1), 183-193 (2001)

JOURNAL
MEDLINE
PUBMED
11415448
REFERENCE
2 (bases 1 to 4150)
AUTHORS
Zhang, L., Ge, L., Tran, T., Stenn, K. and Prouty, S. M.
TITLE
Direct Submission
JOURNAL
Submitted (09-NOV-2000) Skin Biology TRC, Johnson and Johnson,
CPW, 199 Grandview Road, Skillman, NJ 08558, USA

FEATURES
Location/Qualifiers
source
1..4150
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/tissue_type="scalp skin keratinocytes; liver; hair
follicles; brain"
3602..>4150
/gene="SCD"
CAAT_signal
3602..3606
/gene="SCD"
TATA_signal
3869..3872
/gene="SCD"
TATA_signal
3924..3927
/gene="SCD"
mRNA
<3958..>4129
/gene="SCD"
/product="stearyl-CoA desaturase"
/note="3602"
<3958..4129
/gene="SCD"
number=1
3958..4102
/gene="SCD"
4103..>4129
/gene="SCD"
/codon_start=1
/product="stearyl-CoA desaturase"
/protein_id="AAK54510.1"
/db_xref="GI:14150491"
/translation="MPAHLQDD"
4130..>4150
/gene="SCD"
number=1

ORIGIN
Query Match 100.0%; Score 3958; DB 9; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAATTCATCCATTTAAATCATACATAATTAATGGCTTTAGTATATTCACAGGTGTGC 60
DB 1 AGGAATTCATCCATTTAAATCATACATAATTAATGGCTTTAGTATATTCACAGGTGTGC 60

QY 61 ATCCATCAAAATCCATTTTAGAACAGTTTTATTAATCTCCAAATAATAACCTCATTCCTT 120
DB 61 ATCCATCAAAATCCATTTTAGAACAGTTTTATTAATCTCCAAATAATAACCTCATTCCTT 120

QY 121 AGCCATACCCCCCAACATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCAT 180
DB 121 AGCCATACCCCCCAACATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCCAT 180

QY 181 TTTCGTCTCTATAAATTTGGCAATTCGACATTTCAATTAATTAATTAATTAATTAATTAAT 240
DB 181 TTTCGTCTCTATAAATTTGGCAATTCGACATTTCAATTAATTAATTAATTAATTAAT 240

QY 241 GTGAGACTTTGTGACTGGCTGCTTTCACTTAGCATTTCTATTTTAAAGCTCATATGTTA 300
DB 241 GTGAGACTTTGTGACTGGCTGCTTTCACTTAGCATTTCTATTTTAAAGCTCATATGTTA 300

QY 301 CAGTACTTAGCAGTACTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 360
DB 301 CAGTACTTAGCAGTACTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 360

QY 361 CCCATATCATATTTATAGAGACAGGTTCTCATCTGTCTACCCAGGCTGAGTGCAGTGGC 420
DB 361 CCCATATCATATTTATAGAGACAGGTTCTCATCTGTCTACCCAGGCTGAGTGCAGTGGC 420

QY 421 ACAATCATAGTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCTCTACTACCTCAGCC 480
DB 421 ACAATCATAGTCACTGTAACTCAAACTCCCTGGGCTCAAGTGATCTCTACTACCTCAGCC 480

QY 481 TCCAGAGTAGCTAGGACTTAAGGACACACAGGACCATACCTGGCTAAATTTTTTTTAAAT 540
DB 481 TCCAGAGTAGCTAGGACTTAAGGACACACAGGACCATACCTGGCTAAATTTTTTTTAAAT 540

QY 541 TTTTCATTTTATGATTTCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 600
DB 541 TTTTCATTTTATGATTTCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 600

QY 601 ACTTTGTATCCAGGCTGGAGGCGAGTGGCATGTGTGACAGCTGAGAGCCCTTGACTTCCT 660
DB 601 ACTTTGTATCCAGGCTGGAGGCGAGTGGCATGTGTGACAGCTGAGAGCCCTTGACTTCCT 660

QY 661 GGGCTCAAGTGATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720
DB 661 GGGCTCAAGTGATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACAGTGTCA 720

QY 721 CCATGCTGCTGCTGATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 780
DB 721 CCATGCTGCTGCTGATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 780

QY 781 GTACAGTGGCGTAAATATAGTCACTGACGCTCCCTCTCTCTGCGCTCAAGCAATCCGCTG 840
DB 781 GTACAGTGGCGTAAATATAGTCACTGACGCTCCCTCTCTCTGCGCTCAAGCAATCCGCTG 840

QY 841 GCCTCAGCATCTGTAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCAGCCAGCTTAAGTTT 900
DB 841 GCCTCAGCATCTGTAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCAGCCAGCTTAAGTTT 900

QY 901 AAAAAATGATTTTGTGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTATTTG 960
DB 901 AAAAAATGATTTTGTGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATTTTATTTG 960

QY 961 TTGAGACAGGCTCTCACTATGTTGCCATGATCCCCCAGCTCCACTTCCCAAGTGTCA 1020
DB 961 TTGAGACAGGCTCTCACTATGTTGCCATGATCCCCCAGCTCCACTTCCCAAGTGTCA 1020

QY 1021 TCTTATCTGTTTCATTTAGTCAGTTGACAGACATTTAGTGTGTTTCCATTTTGAACATTA 1080
DB 1021 TCTTATCTGTTTCATTTAGTCAGTTGACAGACATTTAGTGTGTTTCCATTTTGAACATTA 1080

QY 1081 TGAATAATACTCCAGTGAATTTCAATGATATACATTTGTGTGGGCATATGTTTTCATTTCT 1140
DB 1081 TGAATAATACTCCAGTGAATTTCAATGATATACATTTGTGTGGGCATATGTTTTCATTTCT 1140

QY 1141 GTTGGTGTATATCTAGGAGTGGAAATCTGATCCCGGGTAAATTTTGAACAGCAGAG 1200
DB 1141 GTTGGTGTATATCTAGGAGTGGAAATCTGATCCCGGGTAAATTTTGAACAGCAGAG 1200

QY 1201 TTCAGGGGAAGAAAAAATTTGGGAAAAATGAAGCATGTTTAGAAAAATCAGCAAGTGCAGGG 1260
DB 1201 TTCAGGGGAAGAAAAAATTTGGGAAAAATGAAGCATGTTTAGAAAAATCAGCAAGTGCAGGG 1260

QY 1261 GTTTTTCGGAGTTTATTTTATTTATTTCTGTTGACAAATGTGCAGTTTGTATGAAGATACAG 1320
DB 1261 GTTTTTCGGAGTTTATTTTATTTATTTCTGTTGACAAATGTGCAGTTTGTATGAAGATACAG 1320

QY 1321 TTATCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380
DB 1321 TTATCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380

QY 1381 AGCACTTTGAATACCAAAATTTAAGGAGCTTGGCTGTAAACAAAAATATAAAAAATACAA 1440
DB 1381 AGCACTTTGAATACCAAAATTTAAGGAGCTTGGCTGTAAACAAAAATATAAAAAATACAA 1440


```

Db      3601  GCCAATGCAACGGCAGGACGAGGTGGCACCAAAATTCCTTCGGCCAAATGACGAGCCGGA 3660
Qy      3661  GTTTACAGAAGCCTCATTTAGCATTTCCCGACAGGCGAGGGCAGAGCCCGGTGG 3720
Db      3661  GTTTACAGAAGCCTCATTTAGCATTTCCCGACAGGCGAGGGCAGAGCCCGGTGG 3720
Qy      3721  TGTGTTGTCGGTGTGCGGAGCATCCCGCGCCCTGCTGCGGTGCGCGAGCCTCGGCC 3780
Db      3721  TGTGTTGTCGGTGTGCGGAGCATCCCGCGCCCTGCTGCGGTGCGCGAGCCTCGGCC 3780
Qy      3781  TCTGTCCTCTCCCTCCCGCCCTTACCTCCACGCGGACCGCGCGCCAGTCAACTCC 3840
Db      3781  TCTGTCCTCTCCCTCCCGCCCTTACCTCCACGCGGACCGCGCGCCAGTCAACTCC 3840
Qy      3841  TCGCACATTTGCCCCCTGCTTGGCAGCGGATAAAGGGGCTCAGGAATATCCGGACACGGT 3900
Db      3841  TCGCACATTTGCCCCCTGCTTGGCAGCGGATAAAGGGGCTCAGGAATATCCGGACACGGT 3900
Qy      3901  CACCGTTGCCAGCTCTAGCTTTAAATTCCTGGGCTCGGGACCTCCACGACCGCGG 3958
Db      3901  CACCGTTGCCAGCTCTAGCTTTAAATTCCTGGGCTCGGGACCTCCACGACCGCGG 3958

RESULT 3
AC018783/c
LOCUS
DEFINITION Homo sapiens clone RP11-2M13, WORKING DRAFT SEQUENCE, 12 unordered
            159351 bp DNA linear HTG 14-MAR-2000
AC018783
VERSION AC018783.3 GI:7229991
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 159351)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Boguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Ferreira,P., FitzHugh,W., Domino,M., Doyle,M., Fenesstor,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,C., Gage,D., Galagan,J.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lechokzy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
and Zody,M.
Direct Submission
Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6692332.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2667
Center clone name: 2_M13
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151339 bases at least Q40
Consensus quality: 155177 bases at least Q30
Consensus quality: 157011 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 158251; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2297: contig of 2297 bp in length
* 2298 2397: gap of 100 bp
* 2398 5719: contig of 3322 bp in length
* 5719 5720: gap of 100 bp
* 5720 10149: contig of 4330 bp in length
* 10149 10249: gap of 100 bp
* 10249 16496: contig of 6247 bp in length
* 16496 16596: gap of 100 bp
* 16596 21823: contig of 5227 bp in length
* 21823 21923: gap of 100 bp
* 21923 25747: contig of 3824 bp in length
* 25747 25847: gap of 100 bp
* 25847 31971: contig of 6124 bp in length
* 31971 32071: gap of 100 bp
* 32071 44449: contig of 12378 bp in length
* 44449 44549: gap of 100 bp
* 44549 64978: contig of 20429 bp in length
* 64978 65079: gap of 100 bp
* 65079 91491: contig of 26413 bp in length
* 91491 91591: gap of 100 bp
* 91591 121937: contig of 30346 bp in length
* 121937 122037: gap of 100 bp
* 122037 159351: contig of 37314 bp in length.
* 122038 159351: contig of 37314 bp in length.
Location/Qualifiers
1. 159351
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-2M13"
/clone_lib="RPC1-11 Human Male BAC"
1. 2297
/note="assembly_fragment"
2398. 5719
/note="assembly_fragment"
5820. 10149
/note="assembly_fragment"
10250. 16496
/note="assembly_fragment"
16597. 21823
/note="assembly_fragment"
21924. 25747
/note="assembly_fragment"
clone_end:SP6
vector_side:right
25848. 31971
/note="assembly_fragment"
clone_end:T7
vector_side:left
32072. 44449
/note="assembly_fragment"
44550. 64978
/note="assembly_fragment"
65079. 91491
/note="assembly_fragment"

```


QY	2043	GGCAGGAGAAATTTCTTGAAOCCGGGAGCAGAGGTTGCAGTGAAGGAGATCACACCCT	2102
DB	57350	GGCAGGAGAAATTTCTTGAAOCCGGGAGCAGAGGTTGCAGTGAAGGAGATCACACCCT	57291
QY	2103	GCACCTCCAGCCTCGGGGAGAGCGGACATTCCTCTCAAAAAACAATAAACAAGAAT	2162
DB	57290	GCACCTCCAGCCTCGGGGAGAGCGGACATTCCTCTCAAAAAACAATAAACAAGAAT	57231
QY	2163	TAAGCAAAATTAGACAATTTGCAGAGAGAACTTGAAGGGGTCAGACCACGTACAGATTTCTG	2222
DB	57230	TAAGCAAAATTAGACAATTTGCAGAGAGAACTTGAAGGGGTCAGACCACGTACAGATTTCTG	57171
QY	2223	TGCCACATGCCAGTACTTCTGTAGGCGATGACTGGATGAGCTGTCCATCTGAAATCATC	2282
DB	57170	TGCCACATGCCAGTACTTCTGTAGGCGATGACTGGATGAGCTGTCCATCTGAAATCATC	57111
QY	2283	CAGTCTTTGTTTCAGAACTTTCACACCGGACAGGAGCCAGGACTGGAAATGCAATCTCTGG	2342
DB	57110	CAGTCTTTGTTTCAGAACTTTCACACCGGACAGGAGCCAGGACTGGAAATGCAATCTCTGG	57051
QY	2343	TCACTGCCACAGAGATTGGCTTTGACCTTCAGACCAGTGGCCAAOAAAGGAGCTGCTTTAG	2402
DB	57050	TCACTGCCACAGAGATTGGCC-TGACCCCTGAGACCAGTGGCCAAOAAAGGAGCTGCTTTAG	56992
QY	2403	TCTACCTCCACGAAATCCOAGTGCTGTCTCTTCCTGGGAAATGAAATCAATTTGGCGCAGCA	2462
DB	56991	TCTACCTCCACGAAATCCOAGTGCTGTCTCTTCCTGGGAAATGAAATCAATTTGGCGCAGCA	56932
QY	2463	CTCGTATTTCTTCCTCTCCACGAGGAGAGATCTCTAGGGCAGTATTTGGGAAAGACATG	2522
DB	56931	CTCGTATTTCTTCCTCTCCACGAGGAGAGATCTCTAGGGCAGTATTTGGGAAAGACATG	56872
QY	2523	GGCATGGAAGGACACCGGGTGAATGCATGCTGCTGCTGTAGCTCTCATGTGAAGG	2582
DB	56871	GGCATGGAAGGACACCGGGTGAATGCATGCTGCTGTAGCTCTCATGTGAAGG	56912
QY	2583	CTCCTTACAGACA CGSAAAAGATGGGGGACAGGACAGATCAGTAGGTCAGAGCATCTC	2642
DB	56811	CTCCTTACAGACA CGSAAAAGATGGGGGACAGGACAGATCAGTAGGTCAGAGCATCTC	56752
QY	2643	AGGACCGAGGGCAATATGTTCTCTGAGCAGGATTAAGAGCTTGGGCTCATATGTGT	2702
DB	56751	AGGACCGAGGGCAATATGTTCTCTGAGCAGGATTAAGAGCTTGGGCTCATATGTGT	56692
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGTCACTACTGTTGCTGTGACCATGGGCAGATTAAT	2762
DB	56691	TTCTGGGCTCAACTGCCAGCTCCATCACTTTACTGTTGCTGTGACCATGGGCAGATTAAT	56632
QY	2763	CCATCTCTCGATATCTCTTTCTCTCACTTTTAAATTGGAAATATGGGTPACCCACTCCCA	2822
DB	56631	CCATCTCTCGATATCTCTTTCTCTCACTTTTAAATTGGAAATATGGGTPACCCACTCCCA	56572
QY	2823	GGGTACAGAGAGGCTTACAGAAAACGATCTTGTGAAATGGCTTGCAGTAAATTAATCAA	2882
DB	56571	GGGTACAGAGAGGCTTACAGAAAACGATCTTGTGAAATGGCTTGCAGTAAATTAATCAA	56512
QY	2883	TACCTGCCAGCTATCTTATTTCCACATCCAGCCCTTTTCGGCTGCTGTGGGTGAAAACA	2942
DB	56511	TACCTGCCAGCTATCTTATTTCCACATCCAGCCCTTTTCGGCTGCTGTGGGTGAAAACA	56452
QY	2943	CATGTCAAGTGTTCCTGACGGTTTTCCAAAGAAGATTTCCAAAATTAACAATGCCAGTC	3002
DB	56451	CATGTCAAGTGTTCCTGACGGTTTTCCAAAGAAGATTTCCAAAATTAACAATGCCAGTC	56392
QY	3003	TGAAGAAATCTCCAAACATCCCGCAGCATCTCTGAGGCGGGGTTGGGGATGGGACTG	3062
DB	56391	TGAAGAAATCTCCAAACATCCCGCAGCATCTCTGAGGCGGGGTTGGGGATGGGACTG	56332
QY	3063	CCGCCCCGGGTCTGAAACAGGATGGTGGCGGAGGCACACACACAGCCAGCCTGTG	3122
DB	56331	CCGCCCCGGGTCTGAAACAGGATGGTGGCGGAGGCACACACACAGCCAGCCTGTG	56272
QY	3123	TGTGGGCGGAGTCCGGGTGGGTCCCGGGTGAACAGCGCGTGGCTGGTGGGCGGGCAG	3182

Db	56271		TTGTCGGCCGGAGTCTCGGTGCGGTCCCGGGTGAGCAGCGGTGGTCTGGTGGCGGGGAG	56212
Qy	3183	AGCATTGTTTCGAGGCGTACCGAGCCCCCGCGCTTCGCCGGGAGGAGGCGGGCTTC	3242	
Db	56211	AGCATTGTTTCGAGGCGTACCGAGCCCCCGCGCTTCGCCGGGAGGAGGCGGGCTTC	56152	
Qy	3243	CCGGT-CCCCAAGCTCCAGATCTCGGGTGGTGGCAGGTCTCCGTGCCACGCGCTGG	3301	
Db	56151	CCGGTCCCCCAAGCTCCAGATCTCGGGTGGTGGCAGGTCTCCGTGCCACGCGCTGG	56092	
Qy	3302	GGGACGGGAAGACGGGACGGAGATCTTAGTGGTGGCGCCCCCGAGGGTTCCACCTG	3361	
Db	56091	GGGACGGGAGACGGGACGGAGATCTTAGTGGTGGCGCCCCCGAGGGTTCCACCTG	56032	
Qy	3362	TTTCTCTGAGAAAATTCCCCAGTCCCAACCCACCGTTCTCGGTGTCGCCGAGGGCGGTC	3421	
Db	56031	TTTCTCTGAGAAAATTCCCCAGTCCCAACCCACCGTTCTCGGTGTCGCCGAGGGCGGTC	55972	
Qy	3422	CTGGCTAGGCTCCGGCGCCGACCCCAACCGGGTCCCGAGCCCTTCCAGAGAAAG	3481	
Db	55971	CTGGCTAGGCTCCGGCGCCGACCCCAACCGGGTCCCGAGCCCTTCCAGAGAAAG	55912	
Qy	3482	CTCCGACGCGGATCCCGGGCAGAGGCCACGCGCGGTGGAGAGAGCTGAGAAGGA	3541	
Db	55911	CTCCGACGCGGATCCCGGGCAGAGGCCACGCGCGGTGGAGAGAGCTGAGAAGGA	55852	
Qy	3542	GAACAAGGGGAGGGGAGCGAGAGCTGGCGGAGAGGGAACAGCAGATTGGCGCGAG	3601	
Db	55851	GAACAAGGGGAGGGGAGCGAGAGCTGGCGGAGAGGGAACAGCAGATTGGCGCGAG	55792	
Qy	3602	CCAAATGGCAACGGCAGGACGAGGTGGCACCAATTCCTTCGGCCATGACGAGCGGAG	3661	
Db	55791	CCAAATGGCAACGGCAGGACGAGGTGGCACCAATTCCTTCGGCCATGACGAGCGGAG	55732	
Qy	3662	TTTACAGAGCCTCATTTAGCATTTTCCCAAGGACAGGGCAGGGGCGAGGCGGGTGGT	3721	
Db	55731	TTTACAGAGCCTCATTTAGCATTTTCCCAAGGACAGGGCAGGGGCGAGGCGGGTGGT	55672	
Qy	3722	GTGGTGTGGGTGTGGCAGCATCCCGCGGCCCTGTCTGCGGTTCGCGAGCCTTCGCGCT	3781	
Db	55671	GTGGTGTGGGTGTGGCAGCATCCCGCGGCCCTGTCTGCGGTTCGCGAGCCTTCGCGCT	55612	
Qy	3782	CTGTCTCTCTCCCGCTCCCGCCTTACTCTCACGCGGGACCGCCCGCGCAGTCAACTCT	3841	
Db	55611	CTGTCTCTCTCCCGCTCCCGCCTTACTCTCACGCGGGACCGCCCGCGCAGTCAACTCT	55552	
Qy	3842	CGCATTGTGCCCCCTCTGTCGACGCGGATAAAGGGGGCTGAGGAAATACCGGACACGGTC	3901	
Db	55551	CGCATTGTGCCCCCTCTGTCGACGCGGATAAAGGGGGCTGAGGAAATACCGGACACGGTC	55492	
Qy	3902	ACCGCTGCGAGCTCTAGGCTTTAAATTCGCCGCTCGGGAGCTTCAGAGCACCGCG	3958	
Db	55491	ACCGCTGCGAGCTCTAGGCTTTAAATTCGCCGCTCGGGAGCTTCAGAGCACCGCG	55435	

RESULT. 4

AL139819

LOCUS

DEFINITION

NOTES

ACCESSION
VERSION

KEYWORDS

SOURCE

ORGANIS

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

requests: clonerequest@sanger.ac.uk
 On Sep 19, 2001 this sequence version replaced gi:14586033.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 RP11-34D15 is from the library RPCI-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-34D15 it may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-34D15 is at 115278 in this
 sequence. The true left end of clone RP11-285F16 is at 56886 in
 this sequence. The true right end of clone RP11-316M21 is at 2000
 in this sequence.

FEATURES

Location/Qualifiers

```

1..115278
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-34D15"
/clone_lib="RPCI-11.1"

```

ORIGIN

```

Query Match      99.0%;   Score 3917.8;   DB 9;   Length 115278;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 3952;   Conservative      0;   Mismatches      2;   Indels      3;   Gaps      3;

QY      3  GAATTCATCCATTTAAATCATACATATTTAATGGCTTTTAGTATATATCATAGTTGTGCAT 62
DB      64592  GAATTCATCCATTTAAATCAATCAATTTAATGGCTTTTAGTATATATCATAGTTGTGCAT 64651

QY      63  CCATCACAAATCCATTTAGAACAGTTTATTTACTCCAAAATAAACCTCGATTCCTTAG 122
DB      64652  CCATCACAAATCCATTTAGAACAGTTTATTTACTCCAAAATTAACCTCGATTCCTTAG 64711

QY      123  CCATCACCCCCCAACATCCTCCATCTCTCTTCCAAAGCCCTGGGCAACCACTACTT 182
DB      64712  CCATCACCCCCCAACATCCTCCATCTCTCTTCCAAAGCCCTGGGCAACCACTACTT 64771

QY      183  TCTGTCCTATAAATTTGCCAATCTGCACATTTTCATATAATGAAGCAACCAATGT 242
DB      64772  TCTGTCCTATAAATTTGCCAATCTGCACATTTTCATATAATGAAGCAACCAATGT 64831

QY      243  GAGACTTTGTGACCTGGCTGCTTCCATTTAGCATTTCTATTTTAAAGGCTCATATGTACA 302
DB      64832  GAGACTTTGTGACCTGGCTGCTTCCATTTAGCATTTCTATTTTAAAGGCTCATATGTACA 64891

QY      303  GTACTTAGCAGTACTCATCTCTTTTATTTTATCTCAATGGTATTCACCTGTGTGGGTATCC 362
DB      64892  GTACTTAGCAGTACTCATCTCTTTTATTTTATCTCAATGGTATTCACCTGTGTGGGTATCC 64951

```

```

QY      363  CATATCATATTTATTAGACAGAGTTCTCTCACTGTGACCCAGGCTGGAGTGCATGGCAC 422
DB      64952  CATATCATATTTATTAGACAGAGTTCTCTCACTGTGACCCAGGCTGGAGTGCATGGCAC 65011

QY      423  AATCATAGCTCACTGTAAACCTCAAACTCCCTGGGCTCAAGTGATCTACTACTCAGGCTC 482
DB      65012  AATCATAGCTCACTGTAAACCTCAAACTCCCTGGGCTCAAGTGATCTACTACTCAGGCTC 65071

QY      483  CAGAGTAGCTAGACATACAGGCAACACAGCCATACCTGGCTAATTTTTTTTTTAATTT 542
DB      65072  CAGAGTAGCTAGACATACAGGCAACACAGCCATACCTGGCTAATTTTTTTTTTAATTT 65131

QY      543  TCATTTTATGTATTCATTTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
DB      65132  TCATTTTATGTATTCATTTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 65191

QY      603  TTGTGTACCAGCTGGAGGCGAGTGGCATGTGTGTGACAGCTGAGCAGCTTGACTTCTCTGG 662
DB      65192  TTGTGTACCAGCTGGAGGCGAGTGGCATGTGTGTGACAGCTGAGCAGCTTGACTTCTCTGG 65251

QY      663  GCTCAAGTGCATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACACAGTGCACC 722
DB      65252  GCTCAAGTGCATCTCTCTGCTCAGCCTCCCAAGTAGCTGGGACTACAAACACAGTGCACC 65311

QY      723  ATGCGCTGGCTGATATTTTTTTTCTTTGAAACAGGGTATCACTCTGTGTGCCAGGCTGGAGT 782
DB      65312  ATGCGCTGGCTGATATTTTTTTTCTTTGAAACAGGGTATCACTCTGTGTGCCAGGCTGGAGT 65371

QY      783  ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCTGGGCTCAAGCAATCCGCTGGC 842
DB      65372  ACAGTGGGCTAATAATAGCTCACTGAGCCTCCCTCTGGGCTCAAGCAATCCGCTGGC 65431

QY      843  CTCAGCATCTCGAGTAGCTGGGACTACAGGCTTTGTGGCCACAGGCCCGAGTAAAGTTTAA 902
DB      65432  CTCAGCATCTCGAGTAGCTGGGACTACAGGCTTTGTGGCCACAGGCCCGAGTAAAGTTTAA 65491

QY      903  AAAATGATTTTGTGTATAGAGAGGCTCTTGCTATCTGCTGCTGAGGCTGTATTTTATCTGT 962
DB      65492  AAAATGATTTTGTGTATAGAGAGGCTCTTGCTATCTGCTGCTGAGGCTGTATTTTATCTGT 65551

QY      963  GAGACAGGCTCTCACTATGTGGCATGATCCCCCACTCCCACTTCCCAAAAGTGTCTATC 1022
DB      65552  GAGACAGGCTCTCACTATGTGGCATGATCCCCCACTCCCACTTCCCAAAAGTGTCTATC 65611

QY      1023  TTATCTGTTCAATAGTCAGTTGACAGATTTAGTGTGTTTCCACTTTTTCACCATATAG 1082
DB      65612  TTATCTGTTCAATAGTCAGTTGACAGATTTAGTGTGTTTCCACTTTTTCACCATATAG 65671

QY      1083  AATAAATCTCCAGTGAATTAATTCATGTATACATTTGTGTGGGCATATGTTTTCATTTCTGT 1142
DB      65672  AATAAATCTCCAGTGAATTAATTCATGTATACATTTGTGTGGGCATATGTTTTCATTTCTGT 65731

QY      1143  TGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTAATATATTTTGACAGGAGATT 1202
DB      65732  TGGGTTTATATCTAGGAGTGAATTTGCTGGATCCCGGTAATATATTTTGACAGGAGATT 65791

QY      1203  CAGGGGAAGAAAACCTTTGGGAAATGAAGCATGTTTGAAGATCAGCAAGAGTGCAGGGGT 1262
DB      65792  CAGGGGAAGAAAACCTTTGGGAAATGAAGCATGTTTGAAGATCAGCAAGAGTGCAGGGGT 65851

QY      1263  TTTTCGAGTTTATTTTATTTTCTGTTGACAAATGTGCAAGTTTGTGATGAAGATCAAGTT 1322
DB      65852  TTTTCGAGTTTATTTTATTTTCTGTTGACAAATGTGCAAGTTTGTGATGAAGATCAAGTT 65911

QY      1323  ATACTAAGTGAGAGTGAGATTAAGGCTGAATAGGGCTTCAAGTAAATCATGAAG 1382
DB      65912  ATACTAAGTGAGAGTGAGATTAAGGCTGAATAGGGCTTCAAGTAAATCATGAAG 65971

QY      1383  CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATTAATAAATAATCACAATT 1442
DB      65972  CACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATTAATAAATAATCACAATT 66031

QY      1443  TTTTTTTTTTTTGTGAAAGAGTCTTTGCTCTTTTCCCTGGCTGGAGGGGAGTGGTGTG 1502

```


Db	66032	TTTTTTTTTTTTTTTGAAAAGAGTCTTGCTCTTTCAOCCCTGGCTGGAGGGCAGTGCTGTG	66091
QY	1503	ATCTCAGCTCACTGCAACTTTGCGCTCCGGGTTCAAGCAATCTCTCCTGCTTCAGCCCTCC	1562
Db	66092	ATCTCAGCTCACTGCACACTTTGCGCTCCGGGTTCAAGCAATCTCTCCTGCTTCAGCCCTCC	66151
QY	1563	CAAGTAGCTGGGACTACAGGCACATCCCAACATGCCAGCTGATTTTGTATTTTGTAGTA	1622
Db	66152	CAAGTAGCTGGGACTACAGGCACATCCCAACATGCCAGCTGATTTTGTATTTTGTAGTA	66211
QY	1623	GAGATGGGATTTTCACTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTGTCATTAATTTGTT	1682
Db	66212	GAGATGGGATTTTCACTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTGTCATTAATTTGTT	66271
QY	1683	GTAACATATTCTTCCTTTTGCTGAGTAGGGCCCCAGACCAAAAAAATAAATCTTAGAA	1742
Db	66272	GTAACATATTCTTCCTTTTGCTGAGTAGGGCCCCAGACCAAAAAAATAAATCTTAGAA	66331
QY	1743	TCCAAATCAGGTGTGGTTTGTACCACTGTCACTTGAGAACCAACAGTGTGACCAAGGCCT	1802
Db	66332	TCCAAATCAGGTGTGGTTTGTACCACTGTCACTTGAGAACCAACAGTGTGACCAAGGCCT	66391
QY	1803	CAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATAATTTCTCCGGGCCAG	1862
Db	66392	CAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATAATTTCTCCGGGCCAG	66451
QY	1863	GGGTGTGGCTCATGCTCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCACCT	1922
Db	66452	GGGTGTGGCTCATGCTCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCACCT	66511
QY	1923	GAGTTCAGAGTTCAAAACACGCTGGCCAAACATGTGAAACCCCGTCTCTACTAAAAAT	1982
Db	66512	GAGTTCAGAGTTCAAAACACGCTGGCCAAACATGTGAAACCCCGTCTCTACTAAAAAT	66571
QY	1983	ACAAAAATTAGCTAAGTCTGTGGCGCATGCTCTGTAATCCCAAGTACTTTGGGAGGGTGA	2042
Db	66572	ACAAAAATTAGCTAAGTCTGTGGCGCATGCTCTGTAATCCCAAGTACTTTGGGAGGGTGA	66631
QY	2043	GGCAGGAGAAATTTCTTGAACCCGGGAGGCAGAGTTGCGAGTGAAGCGAGATCACCACT	2102
Db	66632	GGCAGGAGAAATTTCTTGAACCCGGGAGGCAGAGTTGCGAGTGAAGCGAGATCACCACT	66691
QY	2103	GCATCTCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGAAAT	2162
Db	66692	GCATCTCAGCTGGGGGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGAAAT	66751
QY	2163	TAAGCAATTAGACATTTGACAGAGAGAACTGAAGGGGGTTCAGACCGTACAGATTTCTG	2222
Db	66752	TAAGCAATTAGACATTTGACAGAGAGAACTGAAGGGGGTTCAGACCGTACAGATTTCTG	66811
QY	2223	TGCCACATGCCAAGTACTTTGAGCGCATGATGAGCTGTCCACATCTGAAATCATC	2282
Db	66812	TGCCACATGCCAAGTACTTTGAGCGCATGATGAGCTGTCCACATCTGAAATCATC	66871
QY	2283	CAGTCTTTGTTTCAGAACTTTTCACCGGACAGGGAGCCAGGACTGGAAATGCACTCTCG	2342
Db	66872	CAGTCTTTGTTTCAGAACTTTTCACCGGACAGGGAGCCAGGACTGGAAATGCACTCTCG	66931
QY	2343	TCATGGCCAGAGATTGGCCCTTGAACCTTGAGACCAAGTGGCCAAACAAAGGAGCTGCTTAG	2402
Db	66932	TCATGGCCAGAGATTGGCCCTTGAACCTTGAGACCAAGTGGCCAAACAAAGGAGCTGCTTAG	66990
QY	2403	TCTACCTCCCAAGGAAATCCCAAGGTGCTTGTCTTCTGGGAGTGTGATCATTTGGCGCAGCA	2462
Db	66991	TCTACCTCCCAAGGAAATCCCAAGGTGCTTGTCTTCTGGGAGTGTGATCATTTGGCGCAGCA	67050
QY	2463	CTCCGTATTTTCTCCTCTTCTCCCAAGGGAGGATCCTAGGGCAGTATTTGGGAAAGACATG	2522
Db	67051	CTCCGTATTTTCTCCTCTTCTCCCAAGGGAGGATCCTAGGGCAGTATTTGGGAAAGACATG	67110
QY	2523	GGCATGGAAGACACCGGGTGAATGATAGCTGCTGGTTCTGAGCTCTCATGTGAAGG	2582

Db	67111	GGCATGGAAGGACACACCGGGTGAATGCATAGCCTGCCTGGTTCTGAGCTCTCATGGTAAGG	67117
QY	2583	CTCCTACAGACACGGAAAAAGATGGGGGCAACAGGAGACAGATCAGTAGGGTCAGAGCATCTC	2642
Db	67171	CTCCTACAGACACGGAAAAAGATGGGGGCAACAGGAGACAGATCAGTAGGGTCAGAGCATCTC	67230
QY	2643	AGGACCGAGGGCAATATGTCCTGAGCAGGGATTAGAGCTTGGGCTCTCATATGGTGT	2702
Db	67231	AGGACCGAGGGCAATATGTCCTGAGCAGGGATTAGAGCTTGGGCTCTCATATGGTGT	67290
QY	2703	TTCTGGGCTCAACTGCCAGCTCCGTCACCTTACTTGGTTGCTGTGACCATTGGCAAGTTATT	2762
Db	67291	TTCTGGGCTCAACTGCCAGCTCCATCACTTACTTGGTTGCTGTGACCAATGGCAAGTTATT	67350
QY	2763	CGATCTCTCGATATCTCTTTCCTCACTTTTAAATGGAATAATGGGTACCCACTCCCA	2822
Db	67351	CGATCTCTCGATATCTCTTTCCTCACTTTTAAATGGAATAATGGGTACCCACTCCCA	67410
QY	2823	GGGTACAGAGAGGCTTACAGAAAAAGATCTTGTGATTTGGCTTGCAGTAATTAATTCAA	2882
Db	67411	GGGTACAGAGAGGCTTACAGAAAAAGATCTTGTGATTTGGCTTGCAGTAATTAATTCAA	67470
QY	2883	TACCTGCCAGCTATTCTTATTTCCACATCCAAAGCCCTTTCGCCCTGCTGCTGGTGAAACA	2942
Db	67471	TACCTGCCAGCTATTCTTATTTCCACATCCAAAGCCCTTTCGCCCTGCTGCTGGTGAAACA	67530
QY	2943	CATGTCAGTGTTCCTGACGGTTTCCAAAAAGAAATTCAAAAATTAACAACCTGCCAGTC	3002
Db	67531	CATGTCAGTGTTCCTGACGGTTTCCAAAAAGAAATTCAAAAATTAACAACCTGCCAGTC	67590
QY	3003	TCAGAATCTCCAAACATCCCGCAGCATCTGAGAGGCGGGCTTGGGATGGGACTG	3062
Db	67591	TGAAGAATCTCCAAACATCCCGCAGCATCTGAGAGGCGGGCTTGGGATGGGACTG	67650
QY	3063	CCCGCCGGGTCTCTGAAACAGGATGCGTGGCGCAGGACACACACACAGCCGTGGTGGGGGGCAG	3122
Db	67651	CCCGCCGGGTCTCTGAAACAGGATGCGTGGCGCAGGACACACACACAGCCGTGGTGGGGGGCAG	67710
QY	3123	TGTGGGCGGAGTCCGGTGGGTCGCCGGTTCAGCAGCGCGTGGTGGTGGGGGGGGCAG	3182
Db	67711	TGTGGGCGGAGTCCGGTGGGTCGCCGGTTCAGCAGCGCGTGGTGGTGGGGGGGGCAG	67770
QY	3183	AGCCATTGTCGAGGGCTTACGAGCCCGCCGGCTCGCCCGGAGGAGCGGGGGCTTC	3242
Db	67771	AGCCATTGTCGAGGGCTTACGAGTCCCGCGCTCGCCCGGAGGAGCGGGGGCTTC	67830
QY	3243	CCGGCT-CCCCAAGCTCCAGATCTCTGGGTGGCTGCACGTCTCCCTGCCACGCGCGCTGG	3301
Db	67831	CCGGCTCCCCAAGCTCCAGATCTCTGGGTGGCTGCACGTCTCCCTGCCACGCGCGCTGG	67890
QY	3302	GGGACGGGAAGACGGGACGGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTTCAACACTG	3361
Db	67891	GGGAC- GGAAGA GGGACGGAGATGTTAGTGGTGGGCGCCCCCGAGGGTTTCAACACTG	67949
QY	3362	TTTCTCGAAGTCTCCCGAGTCCCAACCCAGTCTCTCGTGTGCCGAGGGCGCGTTC	3421
Db	67950	TTTCTCGAAGTCTCCCGAGTCCCAACCCAGTCTCTCGTGTGCCGAGGGCGCGTTC	68009
QY	3422	CTGGGCTAGGCTCGCGCCCCCAGCCCCAAAACGGGTCCCGAGCCCTTCACAGAGAGAAG	3481
Db	68010	CTGGGCTAGGCTCGCGCCCCCAGCCCCAAAACGGGTCCCGAGCCCTTCACAGAGAGAAG	68069
QY	3482	CTCCGACGGGGATGCCGGCACAGGCCCCAGCGCGGGTGGAAAGAGTGTGAAGAAGA	3541
Db	68070	CTCCGACGGGGATGCCGGCACAGGCCCCAGCGCGGGTGGAAAGAGTGTGAAGAAGA	68129
QY	3542	GAACAACAGGGAGGGGAGCGAGGACTGCGCGCAGAGGAAACAGCAGATTGCGCCGAG	3601
Db	68130	GAACAACAGGGAGGGGAGCGAGGACTGCGCGCAGAGGAAACAGCAGATTGCGCCGAG	68189
QY	3602	CCAATGGCAACGGCAGGACGAGGTGGCAACAAATTCCTCTCGGCCAAATGACAGGCGGAG	3661
Db	68190	CCAATGGCAACGGCAGGACGAGGACTGCGCGCAGAGGAAACAGCAGATTGCGCCGAG	68249

3662 TTTACAGAAGCCTCATTAGCATTTCOCAGAGGAGGGGCGAGGGGCGAGAGCCGGGTGGT 3721
68250 TTTACAGAAGCCTCATTAGCATTTCOCAGAGGAGGGGCGAGGGGCGAGAGCCGGGTGGT 68309
3722 GTGGTGTGGGTGCGGAGCAGTCCCGCGGCCCTGCTCGGTGCGCGGAGCCTCGGGCCT 3781
68310 GTGGTGTGGGTGCGGAGCAGTCCCGCGGCCCTGCTCGGTGCGCGGAGCCTCGGGCCT 68369
3782 CTGTCTCTCCCTCCCGCCCTTACTCCAGCGGGGACCGCCCGCCGACGTCAACTCCT 3841
68370 CTGTCTCTCCCTCCCGCCCTTACTCCAGCGGGGACCGCCCGCCGACGTCAACTCCT 68429
3842 CGCACTTTGCCCTCTGTTGGCAGCGGATAAAAGGGGCTGAGGAATACCGACACGGTC 3901
68430 CGCACTTTGCCCTCTGTTGGCAGCGGATAAAAGGGGCTGAGGAATACCGACACGGTC 68489
3902 ACCGGTGGCAGCTCAGACCTTAATTCCTCGGCTCGGGGACTCCACGACCGCGG 3958
68490 ACCGGTGGCAGCTCAGACCTTAATTCCTCGGCTCGGGGACTCCACGACCGCGG 68546

RESULT 5
AL391723/c
LOCUS Homo sapiens chromosome 10 clone RP11-38B21, WORKING DRAFT
DEFINITION SEQUENCE, 14 unordered pieces.
ACCESSION AL391723
VERSION AL391723.7 GI:10186824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burton, J.
Direct Submission
Sanger Centre, Hinxton, Cambridgeshire,
Submitted (10-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 19, 2000 this sequence version replaced gi:10178830.

Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

Project Information
Center project name: BA38B21

Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 175002 bases at least Q40
Consensus quality: 179431 bases at least Q30
Consensus quality: 182156 bases at least Q20
Insert size: 184344; sum-of-contigs
Insert size: 140788; 6.0% error; agarose-fp
Quality coverage: 4.15x in Q20 bases; sum-of-contigs Quality
coverage: 6.51x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 59110: contig of 59110 bp in length
* 59111: gap of 100 bp
* 59211: contig of 2075 bp in length
* 61286: gap of 100 bp
* 61386: contig of 2994 bp in length

64380 64479: gap of 100 bp
64480 66747: contig of 2268 bp in length
66748 66847: gap of 100 bp
66848 114355: contig of 47508 bp in length
114356 114455: gap of 100 bp
114456 125406: contig of 10951 bp in length
125407 125506: gap of 100 bp
125507 127673: contig of 2167 bp in length
127674 130111: contig of 2338 bp in length
130112 130212: gap of 100 bp
130213 132354: contig of 2143 bp in length
132355 13454: gap of 100 bp
13455 136537: contig of 4083 bp in length
136538 136637: gap of 100 bp
136638 139127: contig of 2490 bp in length
139128 139227: gap of 100 bp
139228 181355: contig of 42128 bp in length
181356 181455: gap of 100 bp
181456 183524: contig of 2069 bp in length
183525 183624: gap of 100 bp
183625 185644: contig of 2020 bp in length.
Location/Qualifiers
1..185644
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-38B21"
/clone_lib="RPC1-11.1"
1..59110
/note="assembly_fragment:02307"
clone_end:17
vector_side:left
59211..61285
/note="assembly_fragment:00197"
fragment_chain:1
61386..64379
/note="assembly_fragment:01916"
fragment_chain:1
64480..66747
/note="assembly_fragment:00079"
fragment_chain:1
66848..114355
/note="assembly_fragment:00585"
fragment_chain:2
114456..125406
/note="assembly_fragment:02383"
fragment_chain:2
125507..127673
/note="assembly_fragment:00150"
127774..130111
/note="assembly_fragment:00464"
130212..132354
/note="assembly_fragment:00629"
132455..136537
/note="assembly_fragment:00814"
136638..139127
/note="assembly_fragment:00861"
139228..181355
/note="assembly_fragment:02212"
181456..183524
/note="assembly_fragment:02327"
183625..185644
/note="assembly_fragment:02419"

ORIGIN

Query Match 46.2%; Score 1827.2; DB 2; Length 185644;
Best Local Similarity 93.9%; Pred. No. 0; Mismatches 118; Indels 3; Gaps 3;
Matches 1871; Conservative
1968 GTCCTACTAAATAACAAAAATTAGTCTAGTGTGGCGCGATCGCTGTATCCGAC 2027
QY |||||

[illegible]

REFERENCE	1	Enkins, D.K., Winther, M.D., Haardt, M., Goldberg, Y.P., Nwaka, S.O., Ponton, A., Allen, S.J., de Antueno, R.J., and Knickle, L.C.
AUTHORS		Fat regulated genes, uses thereof, and compounds for modulating same
TITLE		Patent: WO 0240666-A 21 23-MAY-2002;
JOURNAL		XENON GENETICS INC (CA)
FEATURES		Location/Qualifiers
source		1. .1970
		/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
ORIGIN		
		Query Match 45.5%; Score 1799.8; DB 6; Length 1970;
		Best Local Similarity 99.7%; Pred. No. 0;
		Matches 1834; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY	2121	GAGAGCGAGCTTCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 2180
DB	1	GAGAGCGAGCTTCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 60
QY	2181	CAGAGAGAACCTGAAGGGGGTTCAGACACGTCACAGATTCTGTGCGCATGCCAAGTACT 2240
DB	61	CAGAGAGAACCTGAAGGGGGTTCAGACACGTCACAGATTCTGTGCGCATGCCAAGTACT 120
QY	2241	TCTGAGGCGATGACTGGATGAGTGTCCACATCTGAAATCATCCAGTCTGTTCAGAACTT 2300
DB	121	TCTGAGGCGATGACTGGATGAGTGTCCACATCTGAAATCATCCAGTCTGTTCAGAACTT 180
QY	2301	TCACACCGGACAGGGAGCAGGACTGGAATGCAGTCTCTGTGCTACTGCGCCAGAGAGTTG 2360
DB	181	TCACACCGGACAGGGAGCAGGACTGGAATGCAGTCTCTGTGCTACTGCGCCAGAGAGTTG 240
QY	2361	GCCTTGACCTGAGACAGTGGCCCAACAAAGAGTGTCTAGTCTACTCTCCAGGAATC 2420
DB	241	CCC-TGACCTGAGACAGTGGCCCAACAAAGAGTGTCTAGTCTACTCTCCAGGAATC 299
QY	2421	CCAGGTGCTGTCTTCTGGGAGTGAATCATGTGGCGAGCACTCGGTATTTCTCTCT 2480
DB	300	CCAGGTGCTGTCTTCTGGGAGTGAATCATGTGGCGAGCACTCGGTATTTCTCTCT 359
QY	2481	TCCACGGGAGAGATCCTAGGGCAGTATTTGGGAAAGACATGGGATGGAAGACACCGG 2540
DB	360	TCCACGGGAGAGATCCTAGGGCAGTATTTGGGAAAGACATGGGATGGAAGACACCGG 419
QY	2541	GTGAATGCATAGCCTGCTGCTGTGAGTCTCATGTGAAGCTCTCATGTGAAGCTCT 2600
DB	420	GTGAATGCATAGCCTGCTGCTGTGAGTCTCATGTGAAGCTCTCATGTGAAGCTCT 479
QY	2601	AGATGGGGGACAGGACAGATCAGTGGGTTCAGAGCATCTCAGGGACCGAGGGCAATAT 2660
DB	480	AGATGGGGGACAGGACAGATCAGTGGGTTCAGAGCATCTCAGGGACCGAGGGCAATAT 539
QY	2661	GGTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGGTGTTCGGGCTCACTGCGCA 2720
DB	540	GGTCTGAGCAGGATTAAGAGCTTGGGCTCTCATATGGTGTTCGGGCTCACTGCGCA 599
QY	2721	GCTCCGCTCACTTACTGTTGCTGACCATGGGCAAGTTATTCATCTCTCCATATCTCT 2780
DB	600	GCTCCATCATCTTACTGTTGCTGACCATGGGCAAGTTATTCATCTCTCCATATCTCT 659
QY	2781	TTCTCTCACTTTAAATGGAATTAATGGGTACCCACCTCCAGGTCACAGAGAGGCTTA 2840
DB	660	TTCTCTCACTTTAAATGGAATTAATGGGTACCCACCTCCAGGTCACAGAGAGGCTTA 719
QY	2841	CAGAAACAGATCTTGTGAATTTGGCTTGCAGTAATTAATCAATACCTGCCAGCTATCTT 2900
DB	720	CAGAAACAGATCTTGTGAATTTGGCTTGCAGTAATTAATCAATACCTGCCAGCTATCTT 779
QY	2901	ATTCACATCCAGCCCTTTGCGCTGCTGCTGGGTGAAAACACATGTCAGTGTTCCTGA 2960
DB	780	ATTCACATCCAGCCCTTTGCGCTGCTGCTGGGTGAAAACACATGTCAGTGTTCCTGA 839

QY	2961	CGGTTTCCACAAAGAGATTTCAAATTTACAACTGCAGTCTGAAGAAATCTCAAAACA 3020
DB	840	CGGTTTCCACAAAGAGATTTCAAATTTACAACTGCAGTCTGAAGAAATCTCAAAACA 899
QY	3021	TCCGCGACGATCTCTGAGCGCGGGCTTGGGATGGGACTGCCCGCCGGGCTCTGAAC 3080
DB	900	TCCGCGACGATCTCTGAGCGCGGGCTTGGGATGGGACTGCCCGCCGGGCTCTGAAC 959
QY	3081	AGGATCGTGTGCGGCGAGGACACACACACAGCAGCCTGTGTGTGCGGCGGAGTCCGG 3140
DB	960	AGGATCGTGTGCGGCGAGGACACACACACAGCAGCCTGTGTGTGCGGCGGAGTCCGG 1019
QY	3141	TGCGGTCTCCGGGTGAGCAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3200
DB	1020	TGCGGTCTCCGGGTGAGCAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079
QY	3201	TACCGAGCCCCCGCGCTTCCCGGAGGAGGCGGGGCTTCCCGCT-CCCCAAGTCC 3259
DB	1080	TACCGAGTCCCCCGCTTCCCGGAGGAGGCGGGGCTTCCCGCTCCCCCAAGTCC 1139
QY	3260	AGATCTCTGGGTGTGCTGCCAGTCTCCCTGCCACGCGCTGGGGGGACGGAACGGGA 3319
DB	1140	AGATCTCTGGGTGTGCTGCCAGTCTCCCTGCCACGCGCTGGGGGGACGGAACGGGA 1198
QY	3320	CGGAGATGTTAGTGTGGGCGCCCCCGGAGGTTCACTACTGTTTCTTGAGAACTTCCC 3379
DB	1199	CGGAGATGTTAGTGTGGGCGCCCCCGGAGGTTTCACTACTGTTTCTTGAGAACTTCCC 1258
QY	3380	CAGTGTCCACCCAGCTTCTCCGTGTGCCCGAGGCGGCTTCTGGGTAGGCTCCGGC 3439
DB	1259	CAGTGTCCACCCAGCTTCTCCGTGTGCCCGAGGCGGCTTCTGGGTAGGCTCCGGC 1318
QY	3440	CCCAGCCCCAACCGGGTCCCAGCCCTTCCAGAGAGAAAGTCCCACGCGGGATGCC 3499
DB	1319	CCCAGCCCCAACCGGGTCCCAGCCCTTCCAGAGAGAAAGTCCCACGCGGGATGCC 1378
QY	3500	GGGCGAGGCCCCAGCGGGGTGGAGAGAGTGTGAGAGAGAGAAACAGAGGGAGGGGG 3559
DB	1379	GGGCGAGGCCCCAGCGGGGTGGAGAGAGTGTGAGAGAGAGAAACAGAGGGAGGGGG 1438
QY	3560	AGCGAGGAGCTGGCGGACAGAGGGAACAGCAGATTGCGCGGAGCCAAATGGCAACGG 3619
DB	1439	AGCGAGGAGCTGGCGGACAGAGGGAACAGCAGATTGCGCGGAGCCAAATGGCAACGG 1498
QY	3620	CGAGTGTGCAACAAATTCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 3679
DB	1499	CGAGTGTGCAACAAATTCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1558
QY	3680	GCATTTCCCGCAGGCGAGGCGGCGAGGCGGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 3739
DB	1559	GCATTTCCCGCAGGCGAGGCGGCGGCGGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1618
QY	3740	GCATTTCCCGCAGGCGAGGCGGCGGCGGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 3799
DB	1619	GCATTTCCCGCAGGCGAGGCGGCGGCGGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1678
QY	3800	GCCCTTACCTCAGCGGGAACCGCGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 3859
DB	1679	GCCCTTACCTCAGCGGGAACCGCGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1738
QY	3860	GGCAGCGGATTAAGGGGGCTGAGGAAATACCGGACACCGTCCACCGCTTCCAGCTCTAG 3919
DB	1739	GGCAGCGGATTAAGGGGGCTGAGGAAATACCGGACACCGTCCACCGCTTCCAGCTCTAG 1798
QY	3920	CTTTTAAATTCGCGGCTCCGGGACCTTCCAGCGCACCGCGG 3958
DB	1799	CTTTTAAATTCGCGGCTCCGGGACCTTCCAGCGCACCGCGG 1837

DEFINITION Papio hamadryas clone RP41-194B17, WORKING DRAFT SEQUENCE.
 AC139668
 VERSION AC139668.1 GI:28273351
 HTG: HTGS PHASE1; HTGS DRAFT.
 KEYWORDS Papio hamadryas (hamadryas baboon)
 SOURCE Papio hamadryas
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Papio.
 REFERENCE 1 (bases 1 to 161827)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 161827)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 COMMENT Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: B051
 Bac Clone Name: RP41-194B17
 This sequence has been compared to sequences of other species using VISTA (<http://www-gsd.lbl.gov/VISTA/>). The results can be viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=ncvalue=SCD
 The order-orientation of the draft sequence was accomplished by using:
 Avid (<http://baboon.math.berkeley.edu/mavid/>),
 Lagan (<http://lagan.stanford.edu/>) and paired end information.
 Funding agent: Programs for Genomic Applications (NHLBI)
 Summary Statistics:
 Sequencing vector: Plasmid; pUC18
 Chemistry: Dye-terminator Big Dye
 Assembly program: Phrap version 0.990329.
 * NOTE: this is a "working draft" sequence. It currently consists of 1 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 161827: contig of 161827 bp in length.
 FEATURES
 source
 Location/Qualifiers
 1..161827
 /organism="Papio hamadryas"
 /mol_type="genomic DNA"
 /db_xref="taxon:9557"
 /clone="RP41-194B17"
 ORIGIN
 Query Match 44.4%; Score 1758; DB 2; Length 161827;
 Best Local Similarity 92.4%; Pred No. 0;
 Matches 1943; Conservative 0; Mismatches 135; Indels 24; Gaps 8;
 QY 1858 GCCAGCGGTGGTCTGATCTGTAATCCAGACATTTGGAGGCGAAGCATGTGGAT 1917
 Db 139308 GCTGGGACGGTGGTCTGATCTGTAATCCAGACATTTGGAGGCGTAAAGGCGCGAT 139367
 QY 1918 CACCTGAGTCTAGGAGTTCAAACCGAGCTGCCACATGTAACCCGCTCTACTA 1977
 Db 139368 CA--TGAGGTGAGGAATGAAGACCGAGCTGACCAATAGGGTGAACCCCGCTGTACTA 139425
 QY 1978 AATAATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTTGGGAG 2037

Db 139426 AATAATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAAATCCAGCTACTTTGGGAG 139485
 QY 2038 GGTGAGGCGAGGAGTAATTTCTTGAAACCCGGGAGCGAGAGTTTCAGTGAAGCGAGATCA 2097
 Db 139486 GGTGAGGCGAGGAGTAATGCTTGAACCCGGGAGCGTGAAGTTTCAGTGAAGCGAGATCA 139545
 QY 2098 CCATCTGCACTCCAGCCTTGGGGAGAGAGCGAGACTTCTTCAAAAAACAAAAACAAA 2157
 Db 139546 CCATCTGCACTCCAGCCTTGGGGAGAGAGCGAGACTTCTTCAAAAAACAAAAACAAA 139605
 QY 2158 AGAATTAAGCAAAATTAGACATTTGCAGAGAGAACTGAAGGGGTGAGACCAAGTACAGAT 2217
 Db 139606 AATAATTCAGCAAAATTAGACATTTGAAGAGAACTGAAGGGGAGAGACCAAGTACAGAT 139665
 QY 2218 TTCTGTGCCACATGCCAAGTACTTCTGAGGCATGACTGATGAGTGTCCACATCTGAAA 2277
 Db 139666 TTCTGTGCCACATGCCAAGTACTTCTGAGGCATGACTGATGAGTGTCCACATCTGAAA 139725
 QY 2278 TCATCCAGTCTTTGTTTCAGAACTTTTCAACCGAGAGGAGCCAGGACTGGAATGCGATCT 2337
 Db 139726 TCATCCAAATCTTTGTTTCAGAACTTTTGCACACTGGACAGGAGCCAGGACTGGTATGCGAGGT 139785
 QY 2338 CCTGGTCACTGGCCAGAGAGTTGGCCCTTGCACCTGAGACCCAGTGGCCAAACAAAGAGAGCTG 2397
 Db 139786 CCTGGTCACTGGCCAGAGAGTTGGCC--TGACCCCTGAGACCCAGTGGCCAAACAAAGAGAGCTG 139844
 QY 2398 CTTAGTCTACCTCCAGGAAATCCAGAGTGTCTTCTCTCTGGGAAAGTGAATCATTTGGGCG 2457
 Db 139845 ATTAGTCTTCTTCCAGGAAATCCAGG-----TGGGAAAGCGAATCATTTAGCG 139892
 QY 2458 CAGCACTCCGTAATTTCTCTCTTCCAGGGAAGGATCTTAGGCGAGTATTTGGGGAAG 2517
 Db 139893 CAGCACTCCGTAATTTCTCTCTTCCAGGGAATGATCTTAGGCGAGATTTGGGCGAG 139952
 QY 2518 ACATGGGCATGGAAGGACACCGGGTGAATGCAATAGCTGCTGCTGTTCTGAGCTCTCATGG 2577
 Db 139953 ACATGGGCATGGAAGGAGAGGGGTGAATGCAATAGCTGCTGCTGTTCTGAGCTCTCATGG 140012
 QY 2578 TAAGGCTCTCTACAGACACGGAAGATGGGGGCACAGGACAGATCATGTAGGTGAGAGC 2637
 Db 140013 TAAGCTCTCTACAGACCCGGAAGAGGGGGGCACAGGACAGATCATGTAGGTGAGAGC 140072
 QY 2638 ATCTCAGGACCCGAGGCAATATGTTCTCAGACGAGGATTAAGAGCTTGGGCTCTCATAT 2697
 Db 140073 ATCTCAGGACCCGAGGACAGTATGTTCTCTGAACAGGATTAAGAGCTTGGGCTCTCATAT 140132
 QY 2698 GGTGTTTCTGGGCTCAATCCAGCTCCGTCACCTTACTGTTGCTGTGACCATGGGCAAG 2757
 Db 140133 GGCATTTCTGGGCTCAATGCGAGCTCCATCATCTTACTGTTGCTGTGATCATGGGCAAG 140192
 QY 2758 TTATTTCCATCTCTCCATATCTTTTCTCTCACTTTTAAATGGAATAATGGGTTGCCACCC 2817
 Db 140193 TTATTTCCATCTCTCTATATATTTTCCACACTTATAAATGGAATAATGGGTTGCCACCC 140252
 QY 2818 TCCAGGGTCAAGAGAGGCTTACAGAAACGATTTCTGTGAATTGGCTTGCAGTAAATAA 2877
 Db 140253 TCCAGGGTCAAGTGAAGGATTAAGAAACGATTTCTGTGAACTGGCTGCGAGTAAATAA 140312
 QY 2878 TTCAATACCTGCCAGCTATTTCTTATTTCCACATCAAGCCCTTTTCCCTGCTGCTGGGTGA 2937
 Db 140313 TTCAATACCTGCCAGCTATTTCTTATTTCCACATCAAGCCCTTTTCCCTGCTGCTGGGTGA 140372
 QY 2938 AAACACATGTCAAGTGTTCCTGACGGTTTCCAAAGAGATTTCCAAATTTACACCTGC 2997
 Db 140373 AAGCACATGTCAAGTGTTCCTGACGTGTTTCCACA--AAGATTTCCAAATTTACACCTGC 140429
 QY 2998 CAGTCTGAAGATCTCCAAACATCCCGACGATCCCTGGAGGCGCGGCTTGGGGATGG 3057
 Db 140430 CAGTCTGAAGATCTCCAAACCTCCCGACGATCCCTGGAGGCGCGGCTTGGGGATGG 140489
 QY 3058 GACTGCCCGCGCGGTCTGAAACAGGATGCGTGGCGGAGGACACACACACACCGAGC 3117

Db	140490	GACTGTTTTCGCCGGGCTCTGAACAGGATTTCGTGCGCGTAGG--CACACACACACACCCAGA	140544
Qy	3118	CTGTGTGTGCGGCCGGAGTCCGGTCCGGTCCCGGAGTCAGACAGCGCTGCGTGTGGTGGCGG	3177
Db	140548	C--TGCCTGCGCGCCGAGTCCGGTTCGGTCCCGGAGTGAACAGCGCTGCTGTGTGGCGG	140605
Qy	3178	GGCAGAGCCATTGTTTCGACAGCGCTACCGAGCCCGCCCGCGCTCGCCCGGAGAGGAGCGCGG	3237
Db	140606	GGCAGAGCCATTGTTTCCAGAGCTGACTGAGCCCCCAGCGCTCGCCCGGAGAGGAGCGCGG	140665
Qy	3238	GCTTCCCGCGTCCCC--AAGCTCCAGATCCTTGGGGTGGTGCACGTCCTCCCTGCGCAGCG	3296
Db	140666	GCTTCCCGCGTCCCGGAGCTCCCGATCCCGGGGTGGCGGCTCGTCTCTCCCTGCGCAGCG	140725
Qy	3297	CCTGGGGGACGGGAGACGGGACGGAGATGTTAGTGTGTGGCGCGCCCCCGCAGGGGTTCCAC	3356
Db	140726	CCTGGGGGACGGGAGAACGGGACGGGAATTGTGTAGTGTGGCGGCTCTCCGAGGCTTCCC	140785
Qy	3357	CACCTGTTTCTTGAGAACTTTCGCCAGTGCACACCCACCCGTTTCTCGTGTGCCCGAGGGC	3416
Db	140786	CACCTGTTTCTTGAGAACTTTCGCCAGTGCACACCCACCCAGCGCTTCGCCCGGGGCG	140845
Qy	3417	CGGCTCTCGGCTAGGCTCCGCGCCCGACGCCCAACCCGGGTCCCAGCGCCCTTCACAGAGA	3476
Db	140846	CGGCTCTCGGCTAGACTCCGCGCCCGACGCCCAACCCGGGTCCCAGTCCCTTCACAGAGA	140905
Qy	3477	GAAAGCTCCGACGCGGGATGCCGGGACAGAGCCCGAGCGGGGTGAAGAACTCTGAG	3536
Db	140906	GAAAGCTCTCTAACCGCGGGGTGCCGGGACAGAGCCCGAGCGGGGTGAAGAACTCTGAG	140965
Qy	3537	AAGGAGAAACAGAGGGGAGGGGAGCGAGAGCTGCGCGACAGAGGAAACAGCAGATTGCG	3596
Db	140966	AAGGAGAAAGAGAGGGGAGGGGAGCGAGAGCTGCGCGACAGAGGAAACAGCAGATTGCG	141025
Qy	3597	CCGAGCCAAATGGCAACCGGACGAGCGAGGTGGCACCAAAATCCCTTCGCGCAATGACGAGC	3656
Db	141026	CCGAGCCAAATGGCAACTGACGAGCAGAGGTGGCACCAAAATCCCTTCGCGCAATGACGAGC	141085
Qy	3657	CGGAGTTTACAGAGCTCATTTAGCAATTTCCCGAGGCGAGGCGAGGGCGAGAGCCCGG	3716
Db	141086	CGGAGTTTACAGAGCTCATTTAGCAATTTCCCGAGGCGAGGCGAGGGCGAGGCGCGCG	141145
Qy	3717	GTGGTGTGTGTGCGGTGTCGGGAGCATCCCGCGCGCCCTGCTGCGGTCCGCGCGAGCTCTC	3776
Db	141146	GTGGTGTGTGTGCGGTGTTGGCAGCAT--CCCGCGCGCCCTGCTGCGGTCTGCGAGCTCT	141204
Qy	3777	GGCCTCTGTCTCTCTCCCTCCCGCCCTTACCTCCAGCGGGAGCCCGCGCGCAGTCAA	3836
Db	141205	GGCCTCTGTCTCTCTCCCTCCCGCCCTTACCTCCAGCGGGAGCCCGCGCGCAGTCAA	141264
Qy	3837	CTCCTCGCACTTTGCCCTCTTTGGCAGCGGATATAAGGGGGCTGAGGAAATACCGGACA	3896
Db	141265	CTCCTCGCACTTTGCCCTCTTTGGCAGCGGATATAAGGGAGCTGAGGAAATACCGGACA	141324
Qy	3897	CGGTCAACCGGTTGCCAGAGCTTAGCTCTTTAAATTCGCGGTTCGGGAGCTTCCAGCACCGC	3956
Db	141325	CGGTCAACCGATGCCAGAGCTTAGCTCTTTAAATTCGCGGTTCGGGAGCTTCCAGCACCGC	141384
Qy	3957	GG 3958	
Db	141385	GG 141386	
RESULT 8			
AC026883			
LOCUS			
DEFINITION			
AC026883 179879 bp DNA linear HTG 29-MAY-2000			
SOURCE			
AC026883			
VERSION			
AC026883.3 GI:8101285			
KEYWORDS			
HTG; HTGS PHASE1; HTGS_DRAFT.			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			

[illegible]

QY 1443 TTTTITTTTTTTTGAGAAAGAGTCTGC 1471
|||||
Dd 111395 TTTTITTTTAAATAAACCGTCATAGTCAGC 111423
|||||

RESULT 9	
AC139669	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

AC139669 190784 bp DNA linear HTG 08-FEB-2003
Lemur catta clone LB9-62P3. WORKING DRAFT SEQUENCE.

ACCESSION AC139669
VERSION AC139669.1 GI:28273352
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Lemur catta (ring-tailed lemur
ORGANISM Lemur catta

Lemur catta (large lemur)
 Lemur catta
 ORGANISM
 Lemur catta
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi
 Mammalia; Eutharia; Primates; Strepsirhini; Lemnridae; Lemur.
 1 (bases 1 to 190784)
 REFERENCE
 Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
 Peng, Z., Malinow, I. and Rubin, E.M.
 AUTHORS

TITLE	Direct Submission	Unpublished
JOURNAL		

REFERENCE
2 (bases 1 to 190784)
Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,
Peng Z., Malinov, I. and Rubin, E.M.

TITLE
JOURNAL
Submitted (08-FEB-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
Direct Submission
August 7, 2003

Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: PGABERK
Center Project Name: L110
Rac Clone Name: LB2-62P23

This sequence has been compared to sequences of other species using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be viewed at:

http://pqa.lbl.gov/cgi-bin/search_cvcqd?type=n&value=SCD

The order-orientation of the draft sequence was accomplished by using:

Avid (<http://baboon.math.berkeley.edu/mavid>),
ragan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>. These libraries are available through the BACPAC Resources Center (<http://thebiologyresource.com/l1/bacpac.htm>) as LBN1-1 to LBN1-4.

Summary Statistics: Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 190784: contra of 190784 bp in length.
* be preserved.

Location/Qualifiers

1. 190784

```
/organism="Lemur catta"
```

/mol type="genomic DNA"

/db xref="taxon:944

ORIGIN

[illegible]

RESULT	10
AY241932	
LOCUS	Bos taurus stearoyl-CoA desaturase variant A (SCD) gene, complete cds.
DEFINITION	
ACCESSION	AY241932
VERSION	AY241932.1 GI:29469126
KEYWORDS	.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

```

REFERENCE
AUTHORS Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
TITLE Genomic sequence of the bovine stearyl-CoA desaturase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 17088)
AUTHORS Medrano,J.F., Islas-Trejo,A.D. and Johnson,A.M.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) Animal Science, University of California,
One Shields Avenue, Davis, CA 95616-8521, USA
FEATURES
source
1..17088
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/chromosome="26"
1..17088
/gene="SCD"
1..1880
/gene="SCD"
1..1880
/gene="SCD"
join(1881..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..17088)
/gene="SCD"
/product="stearyl-CoA desaturase variant A"
join(2025..2051,2568..2850,6720..6850,8314..8519,
10099..10331,13019..13218)
/gene="SCD"
/EC number="1.14.99.5"
/note="delta-9 desaturase"
/codon_start=1
/product="stearyl-CoA desaturase variant A"
/protein_id="AA063569.1"
/db_xref="GI:29469127"
/translation="MPAHLLEIEISSYTTTITAPSPVLONGGKLEKPLYLEE
DIRPEMRDDIYDPTQDEGPKLEYVVRNIIIMSLHLGALYGITLIPCKLYTYI
WLFYLMGALGITAGARHLSHRTYKARLPLRVFLITGNWAFONDFWFSRDRHRAH
KVFSTADPNRSRGFFSHVGMLLVRKHPAVKEKSTLMSDLRAEKLVMFORRY
KPGVLLCFILPTLVPWLWDETFOHSLFFATLFRYALGLNVTWLVNSAAHMYGYRYP
KTINPRENIIIVSLGAGEGPHNYHTFPDYDASEYRWHINFTTFIDCMAAIGLAY
DRKKVSKAAIIRIKRTGEESYKSG"
6926
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
8243
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
8646
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
10153
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
10213
/gene="SCD"
/note="compared to sequence in variant B"
/replaces="c"
10329
/gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
13565
/gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
13638
/gene="SCD"
/note="compared to sequence in variant B"
/replace="c"
13757
/gene="SCD"

```

```

variation
/note="compared to sequence in variant B"
/replace="a"
14047
/gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
14578
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
14662
/gene="SCD"
/note="compared to sequence in variant B"
/replace="y"
15001
/gene="SCD"
/note="compared to sequence in variant B"
/replace="g"
15101
/gene="SCD"
/note="compared to sequence in variant B"
/replace="t"
15202
/gene="SCD"
/note="compared to sequence in variant B"
/replace="a"
15284
/gene="SCD"
/note="compared to sequence in variant B"
/replace="c"
15491
/gene="SCD"
/note="compared to sequence in variant B"
/replace="a"
16524
/gene="SCD"
/note="compared to sequence in variant B"
/replace="t"

```

ORIGIN

```

Query Match      13.2%; Score 521; DB 4; Length 17088;
Best Local Similarity 64.9%; Pred. No. 7.7e-117;
Matches 1232; Conservative 0; Mismatches 490; Indels 177; Gaps 24;

QY 2164 AAGCAATTAGACATTGCAGAGAGAACTGAGGGGGTCAGACACGTCAGATTCTGT 2223
Db 1 AAGCCATTAGACATTGCAG-GAGAGCTGGAGAGGACACACCTGCCAGACTTCT 59

QY 2224 GCC----ACATGCCAAGTACTTCTGAGGCATGACTGGATGAGCTGTCCACATCTGAATC 2279
Db 60 CCCACAAATGCTGAGGGCTTCTGAGCGGTGA-TGAGACAGCTGTCTGTATCTAAAC 118

QY 2280 ATCAGTCTTGTTCAGAACTTTCACCGGACGAGGACGAGCTGAATCAGTCTCC 2339
Db 119 ATCTAGTCTCATTTCAAAACTTTTGCACTGGACAGTGAAGTCAAGTGAAGTGA 178

QY 2340 TGGTCACTGGCCAGAGAGTTGGCCCTTGACCTGAGACCAAGTGGCCCAACAAAGAGTGT 2399
Db 179 TGGTACCGGACGAGGAGCTGGCC-AGACCCCAAGACCTGTGGCCTTCTAGGAGTCT 237

QY 2400 TAGTCTACTCCCGAGAAATCCAGGTGCTTGTCTTCTGGGAAGTGAATCATTTGGCGCA 2459
Db 238 TAATCATCTCCAGGCAATCCAGTGCTCTCTTCC--AGAAGTAGTGTGTGGACACA 295

QY 2460 GCATCCGTAATTTCTCTCTTCCCGGGGAGGATCCTAGGGCAGTATTTGGGAAGAC 2519
Db 296 AAATCTCTCACTTTCTCTCTTCCCA-----AATGATGTATTTGGGAGAC 343

QY 2520 ATGGGCATGGAGGACACCGGGTGAATGATGATACGCTGCTGTGTAGCT----CTCAT 2575
Db 344 AGGGGTGTGTTGGAGCCCGGAGAGATGACAGCCTGCTGTTCCAGCTTATCCCAT 403

QY 2576 GGTAAAGGCTCTTACAGACACGGAAAAGATGGGGGCACAGGACAGATCATGATGGGTGAGA 2635

```

Db 404 GGTAAAGGCTCCACAGCCCGGAAAGGCAAGACTGTGAGCCACAACTAGTGGGTGAGA 463
QY 2636 GCATCTCAGGAGCCGAGGAGCAATATGCTCTGAGCAGGGAATTAAGAGCTTGGGCTCTCAT 2695
Db 464 ACATCTCAGGAGCCAGG-----GATCTGAGCAGGAGGATTAAGA-CTTGGGCTCCATC 514
QY 2696 ATGGTGTCTTCTGGGCTCAACTGCCAGCTCGTCACTTATCTGTTGCTGTGACCATGG-GC 2754
Db 515 ATTACATTTCTGGACTCAGCTACTAGCTCCATCATTTGATTTGGCTGTGTGACCATAAAGC 574
QY 2755 AGTTATTCATCTCTCCATATCTCTTCCTC---ACTTTTAAATGGAATATGGGTA 2811
Db 575 AAGTTATTTAATTTCTCCATAGTCACTCTCTCATCTTTTAAAGGAATAAATGACTA 634
QY 2812 CCACCTCCAGAGGTGACA--GAGAGCTTTACAGAAACGATTTCTTGTGAATTTGGTTCG 2869
Db 635 CTTACCTCAGAGGCAAGATTAAGATCAATGAATGATTTCTTGTGAATTTGGCTTG- 693
QY 2870 AGTAATAAATCAATACCTGCCAGCTATTTATTTTCAATCCAGCCCTTTCCGCTGCTG 2929
Db 694 -----CAATACCTTTTGGCTACTCTTGTGTGCACTTACAGATTAATTTCTCTGCTT 743
QY 2930 CTGGGTGAACAACATCTCAGTGTTCCTC-----TGACGGTTTCCACAAAGAAATTC 2981
Db 744 CTGAGTGAAGACAGCTCTGCTCTTTCTCCAGAAATGGCTTTCCCAACGAGCAATTC 803
QY 2982 CAAATTTACAACTCCAGCTCTGAGAAATCTCCAAACATCCCGCACGCACTCTTGGAGGC 3041
Db 804 CAAACACAGCCGAGCTTCT-----AATCTCCAAACATCCCGCACACCTCTGTTGGC 859
QY 3042 GCGGGCTTGGGATGGAGCTGCCCGCGCGGTC----- 3074
Db 860 ATGGGCAAGGACAGCAGCAGCTCACTGGTGGCTGGGGGTGAACATTTCTTTCGGCGTG 919
QY 3075 -----CTGAAAGAGATGGTGGCGGAGGACACACACACAGCAGCAGCTCTGTGTG 3126
Db 920 ACCATTGGCTGAACCTCTTACGCGGTGCAGATCTACACATCAGCGCATCTACTTGTGTG 979
QY 3127 CGGCGGAGTCCGG-----TGCGGTCCG 3150
Db 980 TGCGGGCTCCAGCTGCGCTGCGGAGGGGAGACCGCAAACTGTTTCAGCAGTCCCG 1039
QY 3151 GGTGAGCAGCGGTGGCTGGTGGCGGAGAGCCATTTGTCGAGCGTACCGA----- 3206
Db 1040 GGTAGCGCTCGCGCTGGTGGCGGCGGCGGCGCATTTGTCGCGCGGAGCAGAACCG 1099
QY 3207 -----GCCCGCGGCTCGCGCGGAGGAGCGGCGGCTTCCC 3244
Db 1100 TCCCTGCGCCCTACCCCGCATCCCGGCGCTCGGACAGAGGAGCGGGGCTTGGC 1159
QY 3245 CGTCCCAAGCTCCAGATCTTGGGTGGTGGCCAGCTCTCGCTGCGCAGCGGCTTGGGG 3304
Db 1160 CGGAGCGCGGCGCGGCGCGGTTTGAAGACAGCTCTCCCTCTCTACGCTCTAGACG 1219
QY 3305 GACGGGAAGACGGGACGAGATTTAGTGGTGGCGGCGCGCGGAGGTTTCACTGTTT 3364
Db 1220 AGCGGAAGGCTGGGGTGGAGTTGTGTGAGTGGTGGCGCTCGTGGTTTCCCACTCTTT 1279
QY 3365 CTTGAGAACTTCCCGAGTCCCAACCCACCTTCTCGTGTGCGCGAGGCGGCTGCTG 3424
Db 1280 CCAGGAAACTTCCCGAGTGGCCATCA-----TTTGGAAATGCGCGGGGCACTCTG 1334
QY 3425 GGCT-----AGGCTCCGCGCCAGCGCCCAACCGGGTCCCGAGCCCTTCCAGAGAAA 3480
Db 1335 GGCTGGCAGCATCCCGGCGCACTCCCGAGTGGGTCTCTCTCTCTCCCGAGCGCT 1394
QY 3481 GCTCCGAGCGGGATGCGGCGAGGCGCCAGCGCGGTGGAGAGAGCTGAGAGG 3540
Db 1395 CAGAGCG--GCAAGGTTGCGGTGAGAGGCGCCAGCGGCCATGTAGAGAGCGG---AGG 1449
QY 3541 AGAAACAGAGGGGAGGGGAGCAGGAGCTGGCGGAGAGGGAACAGCAGATTGGCGCGA 3600

1450 AGAAGGAGGAGGGAGGGGTAGTGAAGAGCTCGCGGACAGAGGAAACAGCAGATTGGCGCGA 1509
QY 3601 GCATATGCAAGCGCAGCAGAGGTGGCACAAATTTCCCTTCGGCAATGACAGAGCCGA 3660
Db 1510 GCCAATGGCAACGGCAGGAGGAGGTGGACCAAAATTTCCCTTCGGCAATGACAGCGCGA 1569
QY 3661 GTTTACAGAAGCCTCATTAGCAATTTCCCCAGAGGAGGCGGCGAGAGGCGGCGGTG 3720
Db 1570 GTCTACAGAAGCC-CATTAGCAATTTCCCGAGGCGAGGCGGAGAGGCGGCGGTGCGGCG 1628
QY 3721 TGTGTTGTCGGGTGTGGGAGCATCCCGGCGGCTCTGCTGGGTGCGCGGAGGCTCGGC 3780
Db 1629 GCCAAGCCGGGTGTGTGTCAGCATTCCTGCTTTCTTTCGGGCCCCAGC--ACGCC 1686
QY 3781 TCTGTCTCTCCCTCCCGCTTACCTCCAGCGGAGCCCGCCGCGCAGTCAACTCC 3840
Db 1687 TCGGCGCTCTGTCTCTCCCTCTCCCGCCATCGGATCTCCACGGTGAACCAACTCT 1746
QY 3841 TCGCACTTTGCCCTCTTGGCAGCGGATAAAAGGGGCTGAGGAAATACCGACAGGT 3900
Db 1747 GCGCACTTTGCCCTTGTGGCAACGAATAAAGAGGTCTGAGGAAATACGGACACAGT 1806
QY 3901 CACCGTTGCCAGCTAGCTTTAAATTCGCCGCTCGG 3939
Db 1807 CACCCCTGCCAGCGCTAGCTTTAAATCCCGACAG 1845

RESULT 11
AX357255 617 bp DNA linear PAT 13-FEB-2002
LOCUS Sequence 1 from Patent WO0162954.
DEFINITION AX357255
ACCESSION AX357255
VERSION AX357255.1 GI:18674409
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Brownlie, A.J., Hayden, M.R., Attie, A.D., Ntambi, J.M., Gray-Keller, M.P. and Miyazaki, M.
TITLE Methods and compositions using streptococcal desaturase to identify triglyceride reducing therapeutic agents
JOURNAL Patent: WO 0162954-A 1 30-AUG-2001; WISCONSIN ALUMNI RESEARCH FOUNDATION
Xenon Genetics Inc. (CA); University of British Columbia (CA)
FEATURES
source Location/Qualifiers
source 1..617
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 12.6%; Score 498.4; DB 6; Length 617;
Best Local Similarity 99.8%; Pred. No. 1.6e-111;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3459 CCCAGCCCTTCCAGAGAGAAAGCTCCCGACGCGGATTCGCGGAGAGGCCAGCGCG 3518
Db 4 CCCCGCCCTTCCAGAGAGAAAGCTCCCGACGCGGATTCGCGGAGAGGCCAGCGCG 63
QY 3519 GGTGGAAGAGAGCTGAGAGAGAAACAGAGGGGAGGGGAGGAGCTGGCGGCGAG 3578
Db 64 GGTGGAAGAGAGCTGAGAGAGAAACAGAGGGGAGGGGAGGAGCTGGCGGCGAG 123
QY 3579 AGGGAACAGCAGATTCGCGCGGAGCAATGCAACGCGGAGGAGGAGGTGGCAAAATTC 3638
Db 124 AGGGAACAGCAGATTCGCGCGGAGCAATGCAACGCGGAGGAGGAGGTGGCAAAATTC 183
QY 3639 CTTTCGCGCAATGACAGCGCGGAGTTTACAGAGCTCATTTAGCATTTCCCGAGGCGAG 3698
Db 184 CTTTCGCGCAATGACAGCGCGGAGTTTACAGAGCTCATTTAGCATTTCCCGAGGCGAG 243

QY	3699	GGCAGGGGCGACAGGC	CCGGGTGGTGGTGTCGGAGCATCCC	CGCGCCCTGCT	3750
DB	244	GGCAGGGGCGACAGGC	CGGGTGGTGGTGTCGGAGCATCCC	CGCGCCCTGCT	303
QY	3759	GCGGTGCCCGCAGACT	CGCGCTCTGTCTCCCTCCGCTCCGCGCCCT	TACTCCACGCGG	3818
DB	304	GCGGTGCCCGCAGACT	CGCGCTCTGTCTCCCTCCGCTCCGCGCCCT	TACTCCACGCGG	363
QY	3819	ACGCGCCCGCCCACT	CAACTCCCTCGCACCTTTGCCCCCTGCTTGGCAGCGGATAAA	AAGGGGG	3878
DB	364	ACGCGCGCGCCAGT	CACCTCCTCGCACCTTTGCCCTGCTTGGCAGCGGATAAA	AAGGGGG	423
QY	3879	CTGAGGAATACCGGA	CACGGTACC CGGTGCCAGCTCTAGCCTTTAAATTC	CGCGCTCG	3938
DB	424	CTGAGGAATACCGGA	CACGGTACC CGGTGCCAGCTCTAGCCTTTAAATTC	CGCGCTCG	483
QY	3939	GGGACCTCCACGAC	CGCGG	3958	
DB	484	GGGACCTCCACGAC	CGCGG	503	
RESULT 12					
AL157935					
LOCUS	AL157935	Human DNA sequence from clone RP11-203J24 on chromosome 9, complete sequence.	164201 bp	DNA linear	PRI 09-JAN-2002
ACCESSION	AL157935				
VERSION	AL157935.28	GI:17221172			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 164201)				
TITLE	Corby,N.				
JOURNAL	Direct Submission				
COMMENT	Submitted (09-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 30, 2001 this sequence version replaced gi:16944853. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM:, EMBL; SW:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr9 RP11-203J24 is from the library RPI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6				
FEATURES	This sequence is the entire insert of clone RP11-203J24 The true left end of clone RP11-379C10 is at 163899 in this sequence. The true right end of clone RP11-228B15 is at 66809 in this sequence.				
SOURCE	Location/Qualifiers 1..164201 /mol_type="genomic DNA"				

Db 5788 CAGCCAGGGGCTGGGATTCGTCGCCAGGTAGGCTGTGCTCTCACATTTTCTTTCTTT 5847
QY 1191 ACAGGCAGAGTTCAGGGGAGAAACTTTGGGAAATAGAGCATGTTTAAATACAGCAA 1250
Db 5848 TCACATTTTCTTTTCTGAGATGGAGTTTCACTCTT---CTGCCAGGCTGGAGTCC 5904
QY 1251 GAGTGCAGGGTTTTTCGGAGTTTATTTTATTTCTGTGTGACAAATGTCAGTTTGATG 1310
Db 5905 GAGACAGCCTCATCAACATGGTGAAACCCCAATGTCGCACTAGCTCACTGTTTCTT 5964
QY 1311 AAGATCAAGTTATATACTAGTGAGAGTGAAGATTAAGGCTGGATAGGCGGTTCAAGT 1370
Db 5965 TTTTTTAGGCTCGT---CAAGTGAAGCAATGAGAGTGGAGAAACAAGCAATCTGTAA 6022
QY 1371 AAATCATGAGACATTTGATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAA 1430
Db 6023 CTAGTTGTGATCAATTAATCTGTGTAACAACAATGTTTTTTCATTTAGTGTGCCATTTAG 6082
QY 1431 AAATCACAAATTTTTTTTTTTTTTTTGGAGAAAGAGTCTTGTCTTTTCAACCCTGGCTGGAG 1490
Db 6083 TTTTTTCTATTTATTTATTTTGGAGATGGAGTCTGTCTGT- TGCCAGGCTGGAA 6141
QY 1491 GGCAGTGGTGTGATCAGCTCATCTGCACTTTCGCTCCCGGTTCAAGCAATTTCTCT 1550
Db 6142 TGCAATGGCGCATCTCGGCTCACTGCACTCCACCTCCCTCTCGGTTCAAGTGAATTCCT 6201
QY 1551 GCTTCACTCCCAAGTAGCTGGAGCTACAGGCACTTCCACCATGCCCCAGCTGATTTT 1610
Db 6202 GCCTTACCTCCCAAGTAGCTGGAGCTACAGGCTGATGCAATGCCACCATGCCAGCTAATTTT 6261
QY 1611 GTATTTTAGTAGAGATGGATTTCACTTTTGTGGCCAAAGCTGGTCTCAAACT- 1663
Db 6262 GTATTTTAGTAGAGATGGATTTTGGCCATGTTTGGCCAGGATGGCTCAAACTCCTGACC 6321
QY 1664 -----TTTTGCTGTCATAATTTGTTGTAATTAATTTGCTTTTGTGAGTAGGAGCCCCCAG 1719
Db 6322 TCAGGTGATCCACCCGCTTGGGCTCCCAAGTCTAGGATTAAGCGGTGAGCCACTGC 6381
QY 1720 ACCAAAAAATAATCTTAGAATCCAAATCAGTGTGTG-----GTTTGACCA 1768
Db 6382 ACCGGCGCTCTCACATTTTTCACACATTTTGTGTAGACATCTTCCCTAGCACAT 6441
QY 1769 CTGTCACTTCAGAACCACTGTGACACAGGCTCAGAGTAGAGTGATCTGTGCTCGA 1828
Db 6442 TTTTCTAACCTTATCCCAAGAGAAAGAAAGAACACAGCAATCTACAAGCCAGTCTATT 6501
QY 1829 AAGAAANTAGATGAAAAATTTCT----- 1853
Db 6502 GAAAGATGACCTCAAGACCTTTCTGAACTAGTGGCATGTGAAACAGGTGCTGAGAC 6561
QY 1854 -----CCGGCCAGCGTGGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCA 1905
Db 6562 TAACACAGGTAGACAGCGGTGGTGGCTTAGGCTGTAATCCCAAGCACTTTGGAGGCTG 6621
QY 1906 AGGCATGTGGATCACTGAGTCAAGAGTTCAAAACAGCTGCCCAACATGTTGAAACC 1965
Db 6622 AGGTGAGCCGATCACTGAGTTGGAGTTCAAGACCAACCTGACCAACGTGGAGAAACC 6681
QY 1966 CCGTCTTCTACTAAAAATACAAAAATTAGCTGAGTGGTGGCCATGCTGTAATCCCA 2025
Db 6682 CCGTCTTCTACTAAAAATACTA- AATTAGCCAGGTGTGGTGGCCGATGCTGTAATCCCA 6739
QY 2026 GCTACTTGGAGGGTGGAGCAGGAGAAATTTCTGAA- CCGGGAGGAGGAGGTTGCAAGT 2084
Db 6740 GCTACTCGGAGGCTGAGCAGGAGAAATCGTTGAAACCCAGGAGGCGGAGGTTGTGGT 6799
QY 2085 AAGCGAGATCAACCACTGCACTCAGCTGG- GGAGAGAGCAGACTTCTCTCAAAA 2143
Db 6800 AGCCAAGATCGCGCATTTGAGCTTAGCTGGCAGCAAGAGTGAATCCGCGCTCAAAA 6859
QY 2144 AAACAAAAACAAGAA 2161
Db 6860 AAAAAAAAAAAAAAAAAA 6877

RESULT 13

AC119725

LOCUS

DEFINITION

AC119725

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC119725 200288 bp DNA linear HTG 27-MAR-2003
Homo sapiens chromosome 3 clone RP11-759D18, WORKING DRAFT
SEQUENCE, 36 unordered pieces.

AC119725.2 GI:29293997
HTG; HTGS PHASE1; HTGS_DRAFT.
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 200288)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbára,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Foster,P., Frantz,P.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Garza,N., Gill,R.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Hamilton,K.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louleseg,H.,
Lozados,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,B., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,K., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczky,R., Woodden,S., Worley,K.,
Wu,C., Wu,Y., Wu,X.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 200288)

Worley,K.C.

Direct Submission

Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 200288)

Worley,K.C.

Direct Submission

Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Mar 27, 2003 this sequence version replaced gi:20376806.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Drafting Center Code: BCM

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HCIR

Center clone name: RP11-759D18

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 169013 bases at least Q40

Consensus quality: 174569 bases at least Q30

Consensus quality: 178019 bases at least Q20

Estimated insert size: 177488; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 36 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1266: contig of 1266 bp in length

1366: gap of unknown length

1367 2448: contig of 1082 bp in length

2449 2548: gap of unknown length

2549 4153: contig of 1605 bp in length

4154 4253: gap of unknown length

4254 5318: contig of 1065 bp in length

5319 5418: gap of unknown length

5419 6987: contig of 1569 bp in length

6988 7087: gap of unknown length

7088 8369: contig of 1282 bp in length

8370 8469: gap of unknown length

8470 9488: contig of 1019 bp in length

9489 9588: gap of unknown length

9589 11671: contig of 2083 bp in length

11672 11771: gap of unknown length

11772 13295: contig of 1524 bp in length

13296 13395: gap of unknown length

13396 14422: contig of 1027 bp in length

14423 14523: gap of unknown length

14523 15982: contig of 1460 bp in length

15983 16082: gap of unknown length

16083 18344: contig of 2262 bp in length

18345 18444: gap of unknown length

18445 20093: contig of 1649 bp in length

20094 20193: gap of unknown length

20194 21347: contig of 1154 bp in length

21348 21447: gap of unknown length

21448 24953: contig of 3506 bp in length

24954 25053: gap of unknown length

25054 27672: contig of 2619 bp in length

27673 27772: gap of unknown length

27773 29866: contig of 2094 bp in length

29867 29966: gap of unknown length

29967 31985: contig of 2019 bp in length

31986 32086: gap of unknown length

32087 34334: contig of 2249 bp in length

34335 37766: contig of 3332 bp in length

37767 37866: gap of unknown length

37867 40539: contig of 2673 bp in length

40540 40639: gap of unknown length

40640 44375: contig of 3736 bp in length

44376 50326: contig of 5851 bp in length

50327 50426: gap of unknown length

50427 53739: contig of 3313 bp in length

53740 53839: gap of unknown length

53840 61829: contig of 7990 bp in length

61830 61929: gap of unknown length

61930 68906: contig of 6977 bp in length

68907 69006: gap of unknown length

69007 78680: contig of 9674 bp in length

78681 78780: gap of unknown length

78781 89431: contig of 10651 bp in length

89432 89531: gap of unknown length

89532 100730: contig of 11199 bp in length

100731 100830: gap of unknown length

100831 112190: contig of 11360 bp in length

112191 112290: gap of unknown length

112291 122606: contig of 10316 bp in length

122607 122706: gap of unknown length

122707 135164: contig of 12458 bp in length

135165 135264: gap of unknown length

135265 147115: contig of 11851 bp in length

147116 147215: gap of unknown length

147216 159369: contig of 12154 bp in length

159370 159469: gap of unknown length

159470 173326: contig of 13857 bp in length

173327 173426: gap of unknown length

173427 200288: contig of 26862 bp in length.

FEATURES

Location/Qualifiers

1..200288

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="3"

/clone="RP11-759D18"

ORIGIN

Query Match 11.3%; Score 447.2; DB 2; Length 200288;

Best Local Similarity 58.5%; Pred.No. 1.6e-98; Indels 111; Gaps 14;

Matches 1099; Conservative 0; Mismatches 668;

QY 365 TATCATATATTAGACACAGGTTCTCACTGTGCACCCAGGCTGGAGTGAGTGGCACA 424

Db 103464 TTTTCTTTCTTTTGGACAGGGTCTTGTCTGTGCACCCAGGCTGGAAACCCAGTGGCACA 103523

QY 425 TCATAGTCTCTGTAAGTCAAACTCTGGCTCAAGTGAATCTTACTACTCAGCCTCCA 484

Db 103524 TCTCAGGTCACTAACCTCTGCTCTGCTGGGTTCAGCAATCTTCTGCTCAGCCTCT 103583

QY 485 GAGTAGCTAGGACTTACAGGCACACACAGCCATCTGGCTAAATTTTTTTTTTAATTC 544

Db 103584 GAGTAGCTAGGACTTACAGGCACACACAGCCATCTGGCTAAATTTTTTTTTTAATTC 103643

QY 545 ATTTTATGTTATTCATTTTCT-----TCTTTTTTTGTTGTTGTTTGTAGATAGG 595

Db 103644 GAGATGGGGTTTCACTATGTGGCCAGGCTGGCTTTTCTTTGTTGTTTGTAGACAGG 103703

QY 596 GTCTCATTGTTTACCCAGGCTGGAGGCGAGTGCATGGTGCACAGCTGA--GCAGCCTTG 653

Db 103704 GTCTTACTCTCTTTTCCAGGCTAGAGTGCATGGTGCATTCATGGCTCAGTGCAGCTCA 103763

QY 654 ACTTCCTGGGCTCAAGTGATCCTCCTCCTCAGCTCCCAAGTAGCTGGGACTACAACA 713

Db 103764 ACCTCTAGGCTCAAGCAATCTCCACCTCGGTTTCCCAAGTGTGGATTAACAGACA 103823

QY 714 CGTGTCAACATGC---CTGGCTGATATTTTTTTTCTTGAACACAGGGATATCATCTGTGTC 770

Db 103824 TGAGCTACCAATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACAGAGTCTTGTCTGTGTC 103883

QY 771 CCAGGCTGGAGTCAAGTGGCGGTATATATAGCTCACTGCAGCCTCCCTCTGGGCTCAAG 830

Db 103884 CCAGGCTGGATTCAGTGGGCAACCTCGGCTCACTGCACTCTGCTCTGCTGTGTCAGG 103943

QY 831 CAATCCGCTGCCCTCAGCATCTCAGTAGCTGGGACTACAGGCTTGTGCCACCCAGGCCCA 890

Db 103944 CAATTCCTCCTCAGCATCTCAGTAGCTGGGACTACAGGCTTGTGCCACCCAGGCCCA 104003

QY	891	GCTAAGTTTTAA	AAAAATGA	TTTTTGGTAT	TAGAGAGGT	CTTGTCTAT	TGTTGCT	CAAGGCTGT	950	
Db	104004	GCTAAATTTTT	-----GTA	TTTTTTAG	TAGAGATGGGGT	TTTTCAGCAT	GTGTGGCCAGAC	TGG	104057	
QY	951	ATTTTTATGTT	GAGACAAGGT	CTCATATG	TGCCATGAT	CCCCCACC	CTCCACAT	TCCC	1010	
Db	104058	TCTCAAACT	CTTTGA-----	-----CCT	CAAAATGAT	CTGCCCGCT	CTGGCCCTCCC	104101		
QY	1011	AAAGTGCTCAT	TTTATCTG	TTCATTAG	TTCAGCAT	TTCACACAGAT	TGTGTTC	CCACTTTT	1070	
Db	104102	AAAGTGCTGGA	TTTATACAGGC	ATGAGCCACT	GCACCCGCCAC	ATTCTTA	TTTTTCACTTA	104161		
QY	1071	TTGACCA	TATGAATA	PATCCAG	TGAATATCAT	GTATACAT	TGTGTGGCGCAT	TGT	1130	
Db	104162	CAGATGAGGGA	ACCAAGAC	CCCTGAGAGGT	GAGTCA	TTTGCTTTGAT	GACACAGGA	T	104221	
QY	1131	TTTTCAT	TTCTGTTGG	TTTATCTAG	GAGTGGAA	TTCCTGGAT	CCCGGTAA	TATTTTG	1190	
Db	104222	CAGCCAGGG	CTGGGATTCG	TGCCCGAG	TAGGCTGT	CTCAAT	TTTTTCTTTCT	TTTT	104288	
QY	1191	ACAGCAGAG	TTTCAGG	GGAGAAAC	CTTTGGG	AAAAATGA	ACATGTTTAG	AAAAATCAGCAA	1250	
Db	104282	TCAC	TTTTTTTTTTT	CTGAGATGGAG	TTTCACTCT	TT-----CCT	GCCAGGCTGGAG	TGC	104338	
QY	1251	GAGTGCAGG	GGTTTTT	TCGGAGTTT	TATTTTAT	TCTGTG	CAAAATGTG	CAGTTTGATG	1310	
Db	104339	GAGACCG	CCCTCATCA	ACATGCTGA	ACCCCAATGGT	CGCATCCT	TAGCTCACT	GTGTTCTT	104398	
QY	1311	AAGATACA	AGTTATCTA	ATGATG	AGAGTGA	GAATTAAG	CGCTGGA	TATAGGCGTT	1370	
Db	104399	TTTTTTAG	GGCTCGT--	CAAGTGA	AGCAATG	ATGAGTGG	AGAGGAACAA	GAAGCAATCTGTA	104456	
QY	1371	AAAATCAT	GAGACCTTTG	GAATACCA	AAATTAAG	GAGGCTTG	CGCTGTAA	ACAAAATAATA	1430	
Db	104457	CTAGTTG	TGATCAATCT	GTAAACA	CAATGTTTTTTTT	CA	TTTTTAGTGTGCC	ATTAG	104516	
QY	1431	AAAATCA	CAATTTTTTTTTTTTT	TTTTTTG	AAAGAGTCTTG	CTTTTCAC	CCCTGGCTGGAG	1490		
Db	104517	TTTTTTTCT	ATTTATTTATTTT	TATTTT	TAGATGGAGTCT	TGCTGT--TGCC	AGGCTGGAA	104575		
QY	1491	GGCAGTGGT	GTGA	TTCTAG	CTCATCTG	CAATTTG	CCCTCCCGGGT	TAAAGCAATCTCCT	1550	
Db	104576	TGCAATGG	CGGATCTCG	GGCTCACT	GCACCTCCA	CTCTCGGGT	TCAAGTATCTCCT	104635		
QY	1551	GCTTCAG	CTCCCAAGT	AGTGGGACT	CAGGCATCTCC	ACATGCC	CCAGCTGATTTTT	1610		
Db	104636	GCCTTAG	CTCCCAAGT	AGTGGGACT	CAGGTGATGCC	ACCATGCC	CAGCTAATTTTT	104695		
QY	1611	GTATTTTT	TAGTAGAGAT	GGGATTT	CAC	TTTGTGGCC	AAAGCTGTGCTCAAACT	1663		
Db	104696	GTATTTTT	TAGTAAAGAT	GGGGTTCC	CCCATGTTGG	CCAGGATGGGCT	CAAACTCCTGACC	104755		
QY	1664	-----TTTT	TGCTGTCA	TAAATTTG	TGTAACTAT	TGTTCC	TTTGTCTGAGGTAG	GGGCCCCCAG	1719	
Db	104756	TCAGGTGAT	CCACCG	CCCTTG	CGCTCCCA	AAAGTGTAG	GAATATAGGCGT	GAGCACATGCTC	104815	
QY	1720	ACCAAAAA	AAAAATAA	ATCTTAGA	ATCCAAATC	AGTGTGTTG-----	-----GTTTGACCA	1768		
Db	104816	ACCGGCG	CGCTCTCACAT	TTTTTCC	CACATTTTT	TGTAG	CACATCTTCCCT	TAGCACAT	104875	
QY	1769	CTGTCA	CTTGAGAAC	CCACAGT	GTGACCG	GGCTCAGG	AGTGTAGGT	GTATCTGCTGCA	1828	
Db	104876	TTTTTCA	CCCTTATCCCA	AGAAGAAA	AGAAACAG	CAATCTA	CAAGCCAGT	GTCTATT	104935	
QY	1829	AAAGAAAT	TAGAAATG	AAAAATTTCT-----	-----	-----	-----	1853		
Db	104936	GAAAGAA	TGACCTCA	AAAGACTTTCT	GAACTAGT	GGCAATG	GAACAGGT	TGCTGAGAC	104995	
QY	1854	-----	-----CCGG	CCAGGCGTGGT	GGCTCATG	CTGTAAT	CCCGAC	ACTTTTGGGAGGCCA	1905	
Db	104996	TAA	GACAGT	TAGAC	CCAGGCGTGGT	GGCTTAC	GCTTAA	TCCCGAC	ACTTTTGGGAGGCTG	105055
QY	1906	AGGCAT	GTGGATCA	CCCTG	AGGTCAGG	AGTTCAA	AAACAG	CTGGCCAA	CAATGGTGA	1965

Db	105056	AGGTGAGCCGATCACTGAGGTTGGAGTTCAAGACCAACTGACCAACGTCGAGAAC	105115
Qy	1966	CCGTCCTTACTAAAAATACAAAAATTAGCTAGTGGTGGCGCATGCTGTATATCCCA	2025
Db	105116	CCGTCCTTACTAAAAATACTA--AATTAGCCAGGTGTGGTGGCGCATGCTGTATATCCCA	105173
Qy	2026	GCTACTTTGGGAGGCTGAGCGCAGGAGAAATTTCTGAA-CCCGGAGGCAGAGTTCGAGTG	2084
Db	105174	GCTACTCTTTTGTCTGAGCGCAGGAGAAATCGCTTGAAACCCCATGAGCGTAGTGTGGTG	105293
Qy	2085	AAGCGAGATCACACACTGCACTCCAGCGCTGGG-GGAGAGAGCGAGACTTCCTCTCAAAA	2143
Db	105234	AGCCAAGATCGCGCAATTGAGCTCTAGCCTGGCGCAGCAAGAGTGAAACTCCGCTCAAAA	105293
Qy	2144	AAACAAAAACAAAAAGAA	2161
Db	105294	AAAAAAAAAAAAAAAAAAAA	105311
RESULT 14			
AL731547			
LOCUS	AL731547	Human DNA sequence from clone RP11-565H13 on chromosome 10,	linear PRI 27-SEP-2002
DEFINITION	AL731547	complete sequence.	
ACCESSION	AL731547	AC073586	
VERSION	AL731547.9	GI:23380932	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 175842)		
AUTHORS	Wray, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,		
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:		
	humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk		
	On Sep 30, 2002 this sequence version replaced gi:21694438.		
COMMENT	Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver		
	Street, Waltham, MA 02453, USA		
	http://www.genomecorp.com		
	----- Genome Center		
	Center: Wellcome Trust Sanger Institute		
	Center code: SC		
	Web site: http://www.sanger.ac.uk		
	Contact: humquerry@sanger.ac.uk		

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>

RP11-565H13 is from the library RPQ1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.		Location/Qualifiers	
FEATURES		source	
		1..175842	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="10"	
		/clone="RP11-565H13"	
		/clone_lib="RP11-11.2"	
ORIGIN			
Query Match		10.9%; Score 433; DB 9; Length 175842;	
Best Local Similarity		58.3%; Pred. No. 4.8e-95;	
Matches 1071; Conservative		0; Mismatches 670; Indels 96; Gaps 14;	
QY	365	TATCATATTATTAGACAGGTTCTCACTCTGTCACCCAGGCTGGAGT---GCAGTGGCA	421
Db	146700	TTTTTTTTTTTTTGACAGATTCTTGCTCTGTCACCCAGGCTGGACTCCGGCAGAGGCA	146759
QY	422	CAATCATAGCTCACTGTAACTCAAACTCTCTGGGCTCAAGTATCTCTACTACTCAGCT	481
Db	146760	CAATCATAGCTCAGTACACCAACAACTCTCTGGGCTCAAGGATCTCTCCACTCAGCT	146819
QY	482	CCAGAGTAGTACGAGCTACAGGCACACAGCCATACCTGGCTAAATTTTTTTTTTAAT	541
Db	146820	CCTGAGTAGCTGGGACTACAGGCACACACACTGCACCTGGCTAATTTTATTTT	146872
QY	542	TTCAATTTATGTTATTCATTTCTTTCTTTTGTGTTGTTGTTTGGATAGGTTCTCA	601
Db	146873	-----TATTTATTTATTTATTTATTTTTCGACACAGAGTCTTG	146910
QY	602	CTTTGTTTACCCAGGCTGGAGGCGAGTGGCATGGTGACAGCTGA--GCAGCTTTGACTTC	659
Db	146911	CTCCATCACCCAGGCTGGAGAGCAGTGGGACACGTTGGCTCACTGCAATCTCTGTCTCC	146970
QY	660	TGGGCTCAAGTGTATCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAAACAGCTGTC	719
Db	146971	TGGGTTCAAGCGAATCTCTCTGCTCAGCTCCCGAGTAGCTGGGACTACAGGCATGTGCC	147030
QY	720	ACCATGGCTGGCTGATATT-----TTTTTTTCTTGAACAGGGTATCACTCTGTGCCCCAG	775
Db	147031	ACCAAGCCTAGCTAAATTTTGTATTTTATAGTAGAGATGGGTTTCTTCCATGTTGGCCAG	147090
QY	776	CTGGAGTACAGTGGCGTAAATATAGCTCACTGCAGGCTCCCTCTGGGCTCAAGCAATC	835
Db	147091	CTGGTCTCGA-----ACTCTGACCTCAGGTGGTCT	147120
QY	836	CGCTGGCTCAGCATCTCTGAGTAGCTGGGACTACAGGCTTTGTGCC--ACCAGGCCAGCT	893
Db	147121	CACCCACCTCGACCTCCCAAGTGTCTGGGATTACAGGCTTGAGCCGGGCCACACCTGGCT	147180
QY	894	AAGTTTAAAAATGATTTTGGTATAGAGGAGTCTTGTCTATGTTGCTCAGGCTGTATT	953
Db	147181	AAATTTTTTAAAAAGAAATTTTGTAGAGA--CCGGCTCTGCTTTGTCAACCAGGTTGGAGT	147239
QY	954	TTTATTTCTTGAGACA-----AGTCTCACTATGTTGCCATGATCCCTCCAGCTCCAC	1005
Db	147240	TCAGTTATGCCATCATAGCTCACTGTAAAGTCGAAATTTCTGGGCTTAAGCAATCTCTCT	147299
QY	1006	TTCCCAAAGTGTCAATCTTATCTGTTTCATTTAGTTCAGTTGACAGACATTTAGTTGTTCC	1065
Db	147300	GGGTCACTTCCAAAAATGCTGGGATTACAGGCGATGAGCCACTGTACTCTGGCCCGAGTTG	147359
QY	1066	ACTTTTGGACCATATTGAATAAATACTCCAGTGAATATTCAATGATATCATTTGTGGGCA	1125
Db	147360	ACTTTCTAGCTCCACTTCTGCAGACTCCCGAGCTTCAAGTAAAGTTCTTTCTCAACTCT	147419
QY	1126	TATGTTTTTCAATTTCTGTTGGTTT-TATATCTAGGAGTGGAAATTTGCTGGATCCCGGTAAT	1184
Db	147420	TTGGTCTTGGACAAAGCAAAACTCTTTCTGCATGGGTCCTTCAAGTCTCTTACCTCAT	147479
QY	1185	ATTTTGTACAGCGACAGTCTCAGGGGAGAAAACTTGGGAAAAATGAAGCATGTTTGAAT	1244
RESULT 15			
AC005037/c			
LOCUS			
DEFINITION			
AC005037			
ACCESSION			
VERSION			
AC005037.2 GI:4827310			
190508 bp DNA linear PRI 30-SEP-2000			
Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.			

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 190508)
MEDLINE Sulston, J.E. and Waterston, R.
PUBMED Toward a complete human genome sequence
99063792 Genome Res. 8 (11), 1097-1108 (1998)
9847074

REFERENCE
AUTHORS 2 (bases 1 to 190508)
TITLE Abbott, A. and Le, T.
JOURNAL The sequence of Homo sapiens BAC clone RP11-469M7
REFERENCE
AUTHORS 3 (bases 1 to 190508)
TITLE Waterston, R.H.
JOURNAL Direct Submission
AUTHORS Submitted (12-JUN-1998) Genome Sequencing Center, Washington
TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
REFERENCE
AUTHORS 4 (bases 1 to 190508)
TITLE Waterston, R.H.
JOURNAL Direct Submission
AUTHORS Submitted (14-MAY-1999) Genome Sequencing Center, Washington
TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
REFERENCE
AUTHORS 5 (bases 1 to 190508)
TITLE Waterston, R.
JOURNAL Direct Submission
AUTHORS Submitted (30-SEP-2000) Department of Genetics, Washington
TITLE University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL On May 14, 1999 this sequence version replaced gi:3309089.
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0469M07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-13J8; the clone sequenced

to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES
source
1. 190508
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-469M7"
/clone_lib="RPCI-11"
155. 222
/rpt_family="L2"
repeat_region
982. 1299
/rpt_family="Alu"
repeat_region
1502. 1533
/rpt_family="(CA)n"
repeat_region
1535. 1823
/rpt_family="Alu"
repeat_region
2005. 2308
/rpt_family="Alu"
repeat_region
2392. 2562
/rpt_family="L1"
repeat_region
3284. 3315
/rpt_family="Alu"
repeat_region
3316. 3338
/rpt_family="(CATA)n"
repeat_region
3339. 3621
/rpt_family="Alu"
repeat_region
3650. 3743
/rpt_family="AT-rich"
repeat_region
4353. 4664
/rpt_family="Alu"
repeat_region
4727. 5018
/rpt_family="Alu"
repeat_region
5545. 5585
/rpt_family="(CATA)n"
repeat_region
5586. 5873
/rpt_family="Alu"
repeat_region
5876. 6007
/rpt_family="Alu"
repeat_region
6327. 6713
/rpt_family="MER21_g"
repeat_region
6924. 7294
/rpt_family="MER2_type"
repeat_region
8027. 8049
/rpt_family="AT-rich"
repeat_region
8050. 8354
/rpt_family="Alu"
repeat_region
8355. 8661
/rpt_family="Alu"
repeat_region
8356. 8684
/rpt_family="AT-rich"
repeat_region
9450. 9752
/rpt_family="Alu"
repeat_region
10371. 10680
/rpt_family="Alu"
repeat_region
10762. 11038
/rpt_family="Alu"
repeat_region
11039. 11109
/rpt_family="Alu"
repeat_region
11138. 11410
/rpt_family="Alu"
repeat_region
11446. 11663
/rpt_family="L1"
repeat_region
11664. 11959
/rpt_family="Alu"
repeat_region
11960. 12189
/rpt_family="L1"
repeat_region
12190. 12334
/rpt_family="Alu"
repeat_region
12335. 12360
/rpt_family="L1"

Db	66518	GGCGGATCTGGCTCTACTGCAAGCTCCGGCTCCGGTTTCAGCCATCTCTCTGCGCTCAA	6645
Qy	848	CATCTCTGATGACTGGGACTACAGCGTTGTGTGCCACAGGCCACAGTAAAGTTTAAATAAT	907
Db	66458	CTTCCCGAGTCTGCTGGGACTACAGTCGCCCGCCACCAAGCTCTGGGTAATTTTTTTTTT	66399
Qy	908	GATTTTTGGTATAGAGGAGTCTTGCTATGTTGGCTCAGGCTGTATTTTATTTATGTTAGAC	967
Db	66398	TATTTTTAGTAGAGACGGGTTTCACTGTGTAGCCAGGATGGTCTCGATCTTCTGACC	66340
Qy	968	AAGGCTCTCATATGTTGGCCATGATCCCCCCCACCTCCACTTCCCAAGTGTCACTTAATC	1027
Db	66339	-----TTTGATCCACCTGCCTCGGCTCCCAAGTGTCTGGGATACA	66297
Qy	1028	TGTTTCATTAGTCAGTTGACAGACATTTAGGTGTTTCCACTTTTTCACCATTTATGAATAA	1087
Db	66296	GGCGTGAAGCCATGACACCAAGCTTAATTTTAAATTTTCTGTAGAGACAGGGTTTCGCTA	66237
Qy	1088	TACTCCAGTGAAATATTCATATGATATACATTTGTGTGGGCATATGTTTTCATTCTGTGGGT	1147
Db	66236	TGTTGCCAGGCTGTGTACTGAA-----CTACTGGCTCAAGCATCTCTGTCTTTGG	66184
Qy	1148	TTATATCTAGGAGTGAAATGCTGGATCCCGGGTAATATTTTGACAGGCAGAGTTTCAGGG	1207
Db	66183	CCACCCAAAGTCTGAGATTACAGGCATGAGCTACATGCTCTGGCTG-----	66136
Qy	1208	GAAGAAAAAATTGGGAAAAATGAAGCATGTTTGAATAATCAGCAAGAGTGCAGGGTTTTTC	1267
Db	66135	-----AATTTTGAATAATAATGACAGTTTAATTTTAAAGTTATATGTAATTTATTATAC	66083
Qy	1268	GGAGTTTATTTATTTATTTCTGTGTGACAAATGTGCAGTTTGATGAGATACAAAGTTATCT	1327
Db	66082	AGCTTAGTTCTCTCGAAAAATCAGCAACTGAAGTGTAAATTTGGTATAGGTGGTAACTCAAGT	66023
Qy	1328	AAGTGAAAGTCTGAGAATTAAGGCTGGAATAGGCGCTTCAGAGTAAAAATCATGAAGCACTT	1387
Db	66022	AGAAAAAGTTAGAGACCAATTAATAAATCTTTATATAAGCAT-----TTT	65975
Qy	1388	TGAATACCAAAATTAAGAGCTTGGCTGTAAACAAATAATAAAAAATCACAATTTTTTTT	1447
Db	65974	TCTGTATAATAACAACATTTTTTCCATACAGGGTAAACCTCTGGGTTTTTTTTTCTGTTGT	65915
Qy	1448	TTTTTTTTTTGAGAAAGAGTCTGCTCTTTTCACTCCCTGGCTGGAGGCGAGTGTGATCTC	1507
Db	65914	TTGTTTTTTGAGATGAGATCTGTTCTGTCAACCGAGCTGGAGTGCATGGCAATCTC	65855
Qy	1508	AGTCTCATGCAACTTTTCGCTCCCGGGTTCAAGCAATTTCTCTGTCTCAGCTCCCAAGT	1567
Db	65854	GGCTCATGCAACTCTGCTCTCCAGGTTCAAGTGATTTCTCTGCTCAGCTCTCTGGAT	65795
Qy	1568	AGCTGGGCTACAGGCACTTCCACCATCCCAAGTCAATTTTTTGTATTTTATGATAGAGA	1626
Db	65794	AGCTTAGACTACAGGACAGTGTCCACACCGCTGGCTCAATTTTCGTAATTTTATGATAGA	65735
Qy	1627	TGGGATTTCACTTTTGTGGCCAGCTGGTCTCAAACCTTTTCTGTCTCAATTTGTGTAA	1686
Db	65734	CGGGTTTACCATGTTAGCCAGGCTGGTCTCAAAC-TCTTGACCTCAAGGGATCTGCT	65676
Qy	1687	CTATTTGTTCTTTGCTGTAGGTAGGCGCCCGCCAGCCAAAAAATAATCTTAGATCCA	1746
Db	65675	GCCTCTCTGCTAGCTTCCCAAAGTCTCGGATTAACAGCGCTGAGCAACCATGCCAGCA	65616
Qy	1747	AATCAGTGTGTGTTTGACCACTGTCACTTGAGAACCAAGTGTGACCGAGGCGCTCAGG	1806
Db	65615	AAACTCTTATTTAT-----AATCAATTCAGATATCTCTAT-----	65579
Qy	1807	AGTAGAGGTGATCTGCTCTGAAAGAGAATAAGATGAAAAATTTCTCCGGCGCAGGCT	1866
Db	65578	-GTGATTTCTCACTTGTACTTTCAAAAATAATTTGTGAAGAAACTTAAGACTATGGCCGGCAT	65520
Qy	1867	GGTGGCTCATGCTGTATATCCCAAGCACTTTGGGAGGCCAAGGCATCTGTGATCACTTGAG	1926
Db	65519	GGTGACTCATGCTGTAAATCCAGCACTTTGGGAAGCCAGGTGGGTGGATCACTTGAG	65460

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:18 ; Search time 2121.36 Seconds

(without alignments)
7926.209 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaatcaccatttaaat.....gggacctccagcaccgcgg 3958

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3958	100.0	4150	ABL57445	Abi57445 Human ste
2	1799.8	45.5	1970	ABK94819	Aah26627 Human ste
3	498.4	12.6	617	AAH26627	Aah26627 Human ste
4	386	9.8	8880	AAK83253	Aak83253 Human imm
5	375.4	9.5	92638	ABQ88096	Abq88096 Human ost
6	368.8	9.3	5491	AAK69044	Aak69044 Human imm
7	366.4	9.3	9742	AAK42015	Aak42015 Genomic s
8	356.2	9.0	84607	AAK90847	Aak90847 Human PAC
9	349.8	8.8	12394	AAH14749	Aad14749 Human gly
10	349	8.8	13216	AAH05122	Aal05122 Human rep
11	349	8.8	13216	AAH05122	Abi98014 Human tes
12	348.6	8.8	396	ABV52505	Abv52505 Human pro
13	347.4	8.8	29163	AAH05121	Aal05121 Human rep
14	347.4	8.8	29163	AAH05121	Abi98013 Human tes
15	347	8.8	6437	AAH42014	Aas42014 Genomic s
16	345	8.7	84495	AAH20588	Aah20588 Human met
17	343.2	8.7	128978	ABK83459	Abk83459 Human cDN
18	343.2	8.7	128978	AAH05121	Aal05121 Human rep
19	342.8	8.7	26928	AAH05121	Aal05121 Human rep
20	342.8	8.7	26928	AAH05121	Abn95780 Gene #227
21	342.8	8.7	26928	AAH05121	Abq77403 Human F2
22	342.8	8.7	26928	AAH05121	Abd88550 Human Fac
23	337.2	8.5	3273	ADB83193	Adb83193 Human cDN

C	24	337	8.5	109906	6	ABK94411	DNA encod
	25	335.2	8.5	128978	6	ABK94459	Human cDN
	26	335.2	8.5	128978	7	AAH54587	Human LIM
C	27	333.8	8.4	183610	7	ACF62736	Cancer ba
	28	333.8	8.4	183610	7	ADB20851	MRP1 base
C	29	333.8	8.4	183610	9	ADB87940	Human UGT
	30	333.8	8.4	183610	9	ADB87940	Human MDR
C	31	333.8	8.4	183610	9	ADB92114	Human GPC
	32	333.6	8.4	35425	9	ADC87616	Human GPC
	33	332	8.4	25715	4	AAH33462	DNA encod
C	34	329.6	8.3	9372	4	AAH36684	Human mus
	35	329.6	8.3	9372	4	AAH36684	Human mus
C	36	329.6	8.3	9372	7	ABX59672	CDNA enco
	37	329.6	8.3	9372	7	ABX59672	CDNA enco
C	38	328.6	8.3	35425	9	ADC87616	Human GPC
	39	328.4	8.3	96593	9	ADC85367	Mouse Blm
	40	328.4	8.3	96594	8	ADA02888	Human BLM
	41	328.4	8.3	96594	9	ADB72626	Human BLM
C	42	328	8.3	8308	4	AAK79820	Human imm
	43	328	8.3	8308	4	AAK68091	Human imm
C	44	327.8	8.3	68355	7	ACF62737	Cancer ba
	45	327.8	8.3	68355	7	ADB20852	MRP1 base

ALIGNMENTS

RESULT 1	ABL57445	ABL57445 standard; DNA; 4150 BP.
ID	ABL57445	
XX	XX	
AC	ABL57445;	
XX	XX	
DI	22-AUG-2002	(first entry)
XX	XX	
DE	Human stearoyl-CoA desaturase gene 5' region.	
XX	XX	
KW	Stearoyl-CoA desaturase; SCD; enzyme; human; promoter; virucide;	
KW	dermatological; cytostatic; immunosuppressive; anti-allergic;	
KW	anti-atheritic; anti-inflammatory; cardiovascular; antianaemic;	
KW	chromosome 10; gene therapy; gene; ds.	
XX	XX	
OS	Homo sapiens.	
PH	Key	Location/Qualifiers
FT	misc_feature	1..3958
FT	misc_feature	/tag= a
FT	misc_feature	/note= "region specifically described in Claim 1"
FT	misc_feature	1..3910
FT	misc_feature	/tag= b
FT	misc_feature	/note= "region specifically described in Claim 2"
FT	misc_feature	complement(105..108)
FT	misc_feature	/tag= c
FT	misc_feature	/bound_moiety= "HNF-3/Fkh-1"
FT	misc_feature	126..129
FT	misc_feature	/tag= d
FT	misc_feature	/bound_moiety= "GREBP"
FT	misc_feature	172..175
FT	misc_feature	/tag= e
FT	misc_feature	/bound_moiety= "NF-Y"
FT	misc_feature	complement(255..258)
FT	misc_feature	/tag= f
FT	misc_feature	/bound_moiety= "NF-Y"
FT	misc_feature	complement(263..266)
FT	misc_feature	/tag= g
FT	misc_feature	/bound_moiety= "IRF-2"
FT	misc_feature	complement(461..464)
FT	misc_feature	/tag= h
FT	misc_feature	/bound_moiety= "GREBP"
FT	misc_feature	complement(541..544)
FT	misc_feature	/tag= i
FT	misc_feature	/bound_moiety= "IRF-1"
FT	misc_feature	complement(669..672)

FT	protein_bind	/tag= j	bound moiety= "SREBP"	FT	protein_bind	bound moiety= "RAR-alpha"
FT	protein_bind	991. .994		FT	protein_bind	2359. .2362
FT	protein_bind	/tag= k		FT	protein_bind	/tag= ai
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	complement(1037. .1040)		FT	protein_bind	complement(2365. .2368)
FT	protein_bind	/tag= l		FT	protein_bind	/tag= aj
FT	protein_bind	bound moiety= "AP-1"		FT	protein_bind	bound moiety= "ER"
FT	protein_bind	complement(1073. .1076)		FT	protein_bind	complement(2382. .2385)
FT	protein_bind	/tag= m		FT	protein_bind	/tag= ak
FT	protein_bind	bound moiety= "RAR-alpha"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	1240. .1243		FT	protein_bind	complement(2414. .2417)
FT	protein_bind	/tag= n		FT	protein_bind	/tag= al
FT	protein_bind	bound moiety= "C/EBP-beta"		FT	protein_bind	bound moiety= "Rel"
FT	protein_bind	1410. .1413		FT	protein_bind	2445. .2450
FT	protein_bind	/tag= o		FT	protein_bind	/tag= am
FT	protein_bind	bound moiety= "NF-1"		FT	protein_bind	bound moiety= "AP-1"
FT	protein_bind	complement(1423. .1426)		FT	protein_bind	complement(2452. .2455)
FT	protein_bind	/tag= p		FT	protein_bind	/tag= an
FT	protein_bind	bound moiety= "HNF3-beta"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	complement(1500. .1503)		FT	protein_bind	complement(2533. .2536)
FT	protein_bind	/tag= q		FT	protein_bind	/tag= ao
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "GR"
FT	protein_bind	complement(1587. .1590)		FT	protein_bind	complement(2726. .2729)
FT	protein_bind	/tag= r		FT	protein_bind	/tag= ap
FT	protein_bind	bound moiety= "NF-kB"		FT	protein_bind	bound moiety= "CREB"
FT	protein_bind	1628. .1631		FT	protein_bind	complement(2744. .2747)
FT	protein_bind	/tag= s		FT	protein_bind	/tag= aq
FT	protein_bind	bound moiety= "NF-kB"		FT	protein_bind	bound moiety= "ER"
FT	protein_bind	complement(1632. .1634)		FT	protein_bind	2862. .2865
FT	protein_bind	/tag= t		FT	protein_bind	/tag= ar
FT	protein_bind	bound moiety= "IRF-1"		FT	protein_bind	bound moiety= "NF-1"
FT	protein_bind	1713. .1716		FT	protein_bind	complement(2917. .2920)
FT	protein_bind	/tag= u		FT	protein_bind	/tag= as
FT	protein_bind	bound moiety= "AP-2"		FT	protein_bind	bound moiety= "PPAR-alpha"
FT	protein_bind	complement(1729. .1732)		FT	protein_bind	complement(2942. .2945)
FT	protein_bind	/tag= v		FT	protein_bind	/tag= at
FT	protein_bind	bound moiety= "HNF3-beta"		FT	protein_bind	bound moiety= "Oct-1"
FT	protein_bind	1833. .1836		FT	protein_bind	2958. .2961
FT	protein_bind	/tag= w		FT	protein_bind	/tag= au
FT	protein_bind	bound moiety= "IRF-2"		FT	protein_bind	bound moiety= "CREB/ATF"
FT	protein_bind	complement(1863. .1866)		FT	protein_bind	complement(2965. .2968)
FT	protein_bind	/tag= x		FT	protein_bind	/tag= av
FT	protein_bind	bound moiety= "Whn"		FT	protein_bind	bound moiety= "STAT"
FT	protein_bind	1917. .1920		FT	protein_bind	complement(2979. .2982)
FT	protein_bind	/tag= y		FT	protein_bind	/tag= aw
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "C/EBP-beta"
FT	protein_bind	1925. .1928		FT	protein_bind	3056. .3059
FT	protein_bind	/tag= z		FT	protein_bind	/tag= ax
FT	protein_bind	bound moiety= "T3R, RAR-alpha"		FT	protein_bind	bound moiety= "NF-kB"
FT	protein_bind	complement(2087. .2090)		FT	protein_bind	complement(3062. .3066)
FT	protein_bind	/tag= aa		FT	protein_bind	/tag= ay
FT	protein_bind	bound moiety= "Pax-6"		FT	protein_bind	bound moiety= "NF-kB"
FT	protein_bind	2093. .2096		FT	protein_bind	complement(3065. .3068)
FT	protein_bind	/tag= ab		FT	protein_bind	/tag= az
FT	protein_bind	bound moiety= "SREBP"		FT	protein_bind	bound moiety= "SP-1"
FT	protein_bind	complement(2116. .2119)		FT	protein_bind	3131. .3134
FT	protein_bind	/tag= ac		FT	protein_bind	/tag= ba
FT	protein_bind	bound moiety= "AP-2"		FT	protein_bind	bound moiety= "Ets-1"
FT	protein_bind	complement(2179. .2182)		FT	protein_bind	complement(3161. .3164)
FT	protein_bind	/tag= ad		FT	protein_bind	/tag= bb
FT	protein_bind	bound moiety= "CHOP"		FT	protein_bind	bound moiety= "Whn"
FT	protein_bind	2199. .2202		FT	protein_bind	complement(3162. .3166)
FT	protein_bind	/tag= ae		FT	protein_bind	/tag= bc
FT	protein_bind	bound moiety= "RAR-alpha"		FT	protein_bind	bound moiety= "Myc/Max"
FT	protein_bind	complement(2231. .2234)		FT	protein_bind	3173. .3176
FT	protein_bind	/tag= af		FT	protein_bind	/tag= bd
FT	protein_bind	bound moiety= "NF-1"		FT	protein_bind	bound moiety= "SP-1"
FT	protein_bind	2245. .2248		FT	protein_bind	32

KW dyslipidaemia; atherosclerosis; coronary artery disease;
KW cerebrovascular disease; peripheral vascular disease; inflammation;
KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;
KW acne; body weight disorder; obesity; cachexia; anorexia;
KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;
KW diabetes; diabetic complication; genetic polymorphism; da.
XX Homo sapiens.
XX
XX W0200240666-82.
XX
XX 23-MAY-2002.
XX
XX 19-NOV-2001; 2001WO-CA001632.
XX
XX 17-NOV-2000; 2000US-0248589P.
XX
XX (XENO-) XENON GENETICS INC.
XX
XX Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ;
XX Ponton A, De Antueno RJ, Jenkins DK, Nwaka SO;
XX
XX WPI; 2002-508327/54.
XX
XX Novel isolated polypeptide segment encoded by fat regulated genes, useful
XX for diagnosing the presence of or a predisposition for a disorder
XX involving fatty acid regulated genes in a subject.
XX
XX Claim 23; Fig 26; 225pp; English.
XX
XX The invention describes an isolated polypeptide segment (I) whose genes
XX are fat regulated. (I) or the polynucleotide encoding it (II) are useful
XX for diagnosing the presence of or a predisposition for a disorder
XX involving fatty acid regulated genes in a subject. A composition
XX containing (I) or (II) is useful for treating a disorder involving fatty
XX acid regulated genes, where the disorder is selected from a
XX polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular
XX disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,
XX coronary artery disease, cerebrovascular disease or peripheral vascular
XX disease), inflammation (such as sinusitis, asthma, pancreatitis,
XX osteoarthritis, rheumatoid arthritis or acne), body weight disorders
XX (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,
XX cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic
XX complications. (I) or (II) is useful as research agent and materials for
XX discovery of treatments and diagnostics for a disease, particularly human
XX disease. (II) is useful for constructing nucleotide probes and primers,
XX for detecting genetic polymorphism, for detecting changes in the level of
XX expression of (II), and as a diagnostic tool. This sequence represents a
XX control region from a gene encoding a fatty acid regulated protein
XX
XX Sequence 1970 BP; 428 A; 590 C; 588 G; 364 T; 0 U; 0 Other;
XX
XX Query Match 45.5%; Score 1799.8; DB 6; Length 1970;
XX Best Local Similarity 99.7%; Pred. No. 0;
XX Matches 1834; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
XX
XX 2121 CAGAGCGAGACTTCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 2180
XX 1 GAGAGCGAGACTTCTCTCAAAAAAACAACAAAGAAATTAAGCAAAATAGACATTG 60
XX
XX 2181 CAGAGAGAACTGAAGGGGGTCAGACCAAGTACAGATTCTGTGCCAATGCCAAGTACT 2240
XX 61 CAGAGAGAACTGAAGGGGGTCAGACCAAGTACAGATTCTGTGCCAATGCCAAGTACT 120
XX
XX 2241 TCTGAGGATGACTGGATGAGTGTCCACATCTGAAATCATCCAGTCTGTGTGCAACTT 2300
XX 121 TCTGAGGATGACTGGATGAGTGTCCACATCTGAAATCATCCAGTCTGTGTGCAACTT 180
XX
XX 2301 TCACACCGGACAGGGAGCAGGACTGGAAATGAGTCTCCTGTGCTAGTGGCCAGAGAGTTG 2360
XX 181 TCACACCGGACAGGGAGCAGGACTGGAAATGAGTCTCCTGTGCTAGTGGCCAGAGAGTTG 240
XX
XX 2361 GCCTTGACCTGAGACCAAGTGGCCAAACAAGAGAGTGTCTTACCTCCCAAGAAATC 2420

Db 241 GCC-TGACCCCTGAGACAGTGGCCCAACAAGAGAGTCTTGTAGTCTACCTCCAGGAATC 299
Qy 2421 CCAGGTGCTGCTCTCTGAGAGTGAATCATTTGGCGCAGCACTCCGATTTTCTCTCT 2480
Db 300 CCAGGTGCTGCTCTCTGAGAGTGAATCATTTGGCGCAGCACTCCGATTTTCTCTCT 359
Qy 2481 TCCAGGGGAGAGATCCTAGGGGAGATTTTGGGAAAGACATGGGCATGGAAGGACACCGG 2540
Db 360 TCCAGGGGAGAGATCCTAGGGGAGATTTTGGGAAAGACATGGGCATGGAAGGACACCGG 419
Qy 2541 GTGAATGCATAGCCTGCTCTGAGCTCTCATGTGTAAGGCTCTTACACACACGGA 2600
Db 420 GTGAATGCATAGCCTGCTCTGAGCTCTCATGTGTAAGGCTCTTACACACACGGA 479
Qy 2601 AGATGGGGGACACGGGACAGATCAGTAGGGTTCAGAGCATCTCAGGGACCGAGGCAATAT 2660
Db 480 AGATGGGGGACACGGGACAGATCAGTAGGGTTCAGAGCATCTCAGGGACCGAGGCAATAT 539
Qy 2661 GGTCTCTGAGCAGGAGATTAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 2720
Db 540 GGTCTCTGAGCAGGAGATTAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 599
Qy 2721 GGTCTCTGAGCAGGAGATTAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 2780
Db 600 GGTCTCTGAGCAGGAGATTAAGAGCTTGGGCTCTCATATGGTGTCTTCTGGGCTCAATGCGCA 659
Qy 2781 TTCTCTCACTTTTAAATGGGATTAATGGGATACCCACCTCCAGGGTTCAGAGAGGCTTA 2840
Db 660 TTCTCTCACTTTTAAATGGGATTAATGGGATACCCACCTCCAGGGTTCAGAGAGGCTTA 719
Qy 2841 CAGAAAAAGATTTCTGGAATTTGGCTTGCAGTAATAATCAATACCTGCCAGCTATTCTT 2900
Db 720 CAGAAAAAGATTTCTGGAATTTGGCTTGCAGTAATAATCAATACCTGCCAGCTATTCTT 779
Qy 2901 ATTCCACATCAAGCCCTTTCGCTGCTGTGGTGTGAAACACATGTCAGTGTTCCTGA 2960
Db 780 ATTCCACATCAAGCCCTTTCGCTGCTGTGGTGTGAAACACATGTCAGTGTTCCTGA 839
Qy 2961 CGGTTCACAAAGAAATTTCCAAATTTACAACTGCGCAGTCTGAAGAAATCTCAAAACA 3020
Db 840 CGGTTCACAAAGAAATTTCCAAATTTACAACTGCGCAGTCTGAAGAAATCTCAAAACA 899
Qy 3021 TCCCGCAGCATCTCTGAGGCGGGCTTGGGATGGAGCTGCGCCGCGCGGTCTCTGAAC 3080
Db 900 TCCCGCAGCATCTCTGAGGCGGGCTTGGGATGGAGCTGCGCCGCGCGGTCTCTGAAC 959
Qy 3081 AGATGCGTGGCGCAGGACACACACACAGCAGCTGTGTGTGGCGCGGAGTCCGG 3140
Db 960 AGATGCGTGGCGCAGGACACACACAGCAGCTGTGTGTGGCGCGGAGTCCGG 1019
Qy 3141 TGCGGTCCCGGTGAGCAGCGGTGGTGGGCGGGGAGAGCCATTGTCGAGGCG 3200
Db 1020 TGCGGTCCCGGTGAGCAGCGGTGGTGGGCGGGGAGAGCCATTGTCGAGGCG 1079
Qy 3201 TACCGAGCCCGCGCTGCGCGGGAGGAGGCGGGCTTCCCGCT-CCCCAAGTCTC 3259
Db 1080 TACCGAGTCCCGCGCTGCGCGGGAGGAGGCGGGCTTCCCGCTCCCGGAGTCC 1139
Qy 3260 AGATCTCGGGTGGCTGCCAGCTCTCCCTGCCAGCGCCCTGGGGGGAGCGGAGACGGGA 3319
Db 1140 AGATCTCGGGTGGCTGCCAGCTCTCCCTGCCAGCGCCCTGGGGGGAGCGGAGACGGGA 1198
Qy 3320 CGGAGATGTTAGTGGTGGCGCGCCCGGAGGTTTACACCTGTTTCTCTGAGAACTTCCC 3379
Db 1199 CGGAGATGTTAGTGGTGGCGCGCCCGGAGGTTTACACCTGTTTCTCTGAGAACTTCCC 1258
Qy 3380 CAGTGGCCACCCAGCCCTTCTCGGTGTGCCGAGGCGGGTCTTGGGTAGGCTCCGCGC 3439
Db 1259 CAGTGGCCACCCAGCCCTTCTCGGTGTGCCGAGGCGGGTCTTGGGTAGGCTCCGCGC 1318
Qy 3440 CCAGGCCCCAAAACCGGTCTCCAGCCCTTCCAGAGAGAAAGCTCCCGACCGGGATGCC 3499

QY	944	AGGCTGATTTTTATTGTTGAGCAAGGTCCTCACTATGTTGCCATGATCCCCCACTCC	1000
Db	7055	ACCATCCCGGCTGCAGTGTATTTATTCCTTTTGTGCTGAATAGTATTCATTTTATGG	7114
QY	1004	ACTTCCCAAGTGTCTATCTATTAATCTGTTCAATAGTCAGTTGACAGACATTTAGTGTGTTT	1063
Db	7115	ACATAC-----TCATATTTTATGCACTTCATTTGTGCAGCTAAATAGACATTTGGTAACCTT	7168
QY	1064	CCACTTTTTCACCAATTAATGAATAATCTCCAGTGAAATATTCATGT---ATACATTTGTGT	1120
Db	7169	CTACTTTTTCGATATTAAGAAATATAGCTTCTCGGAACATTCGTATATCAAGATTTTAATGT	7228
QY	1121	GGGCATATGTTTTCAATTCGTGGGTTTATATCTAGGAGTGGAAATGCTGGATCCCGGG	1180
Db	7229	GGATCTATGTTTTCAITTTTTCTTGGGTATGTACCTAGGAGCAGAAATGATGSGTCAATGTG	7288
QY	1181	TAATATTTTGACAGGCGAGGTTGAGGGGAAGAAAAAATCTCGGAAAAATGAAGCATGTTTAG	1240
Db	7289	GTGACTCTGTAGAAGCTTTTCAAGCTGCGGTACTGTTTCCAGAGTGGCTGCATATTTTCC	7348
QY	1241	AAATCAGCAAGAGTGCAGGGGTTTTTTCGGAGTTTTTATTTTATATTCCTGTTGACAAATGTG	1300
Db	7349	ATTTGAGTTGAAACATCTTTTCATGGGCTTATCGACGATTTTGATATCTTTTGGAGAAATGC	7408
QY	1301	CAGTTTGATGAAGATACAAGTTATCTATAGTGAGAGTGAATTAAGGCTGGAATAGGG	1360
Db	7409	CCATTCGGTTCTTTTGCTGATTTTTTAATTTATTTATCTCTTTTATTAATGAGTTGTAAGCG	7468
QY	1361	CGTTCAGAGTAAAAATCATGAAGCNCITTTGAATACCAAAATTAAGGAGCTTGGCTGTAAAC	1420
Db	7469	AAATCTTTTAGATAGTCCAGATACAAGTCCCTTATCAGA-----TATAT	7511
QY	1421	AAAAATAAAAAAATCAAAATTTTTTTTTTTTTTTTTTTCAGAAAGAGTCTTGCTCTTTCAAC	1480
Db	7512	TATTTGCAATATTTCTCCCAATTTTGTGGGTGTCTTTGCAGCTTCTTGCTTTTTCAA	7571
QY	1481	CTGGCTGGAGGCGAGTGGTGTGATCTCAGCTACTGCAATTTTGCTCTCCCGGTTCAAG	1540
Db	7572	TTG-----CAGTAAAAATATACATAACATAAAATTTAAACATTTTGGGGTTGGGGCAAT	7625
QY	1541	CAATTCCTGCTTCAGCCTCCCNAGTAGCTGGGACTACAGGCNCTTCCACACATGCCCA	1600
Db	7626	GGTTCACACCTGTAAATCCCAACGCAATTTAGAGGCGCAGGTAAGGAGGANTGCTTCAGCCCA	7685
QY	1601	GCTGATTTTTGTATTTTTTAGTAGAGATGGCAATTCACATTTGTTGGCCCAAGCTGCTCTCA	1660
Db	7686	GGAGTTTGACACAGCTCTGGGCAACATAGCAAGATCTGTCTACAAA-----	7734
QY	1661	ACTTTTTGCTGTATAAATGTTGTAACTATTGTTTCCTTTTCTGAGTAGGGGCCCCAGA	1720
Db	7735	-----ATATTTTAAAAATTAGCTGGGCTGGCTGCACACTTCTAGTCCCAGGA	7781
QY	1721	CCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTGTGTTTGACACATGTCACITGAG	1780
Db	7782	TGAGGTGAGAAAAATCGCTTGAGTCCAGGAGGTCAGGGCTCTGAAGTAGGCGGTGATCACAC	7841
QY	1781	AACCAAGCTGTGACAGGGCTCAGGAGTAGAGGTGATCTCTGCTGAAAGAGAAATAGA	1840
Db	7842	CACTGCAAGTCCAGCGCTG-----ATGCGAGATGAGACTCTGTATTTTTTAAAAAATA	7894
QY	1841	ATGAAAAATATTCCTCGGGCCAGGCGTGTGGCTCATGCTGTAAATCCAGCACTTTGGGA	1900
Db	7895	AACAAAAACAAABAGAGGCCAGGCGGGTGGCTCACGCTGTAAATCCAGCACTTTGGGG	7954
QY	1901	GGCCAAAGGCATGTGGATCACTGAGGTCAAGAGTTTCAAAACCGACCTTGGCCCAATGGTG	1960
Db	7955	GACCAAGCGCGGAGATCACTGAGTTCGGGAGTTTCGAGACCGACCTTGACCTACACGGAG	8014
QY	1961	AAACCCGCTCTCTACTAAAAATAAAAAAATTAGCTAAGTGTGTGGCGCATGCTCTGAA	2020
Db	8015	AAACCGTGCCTCAACTGAAATTC-----AAAAATTTCCAGGTGTGGTGTGGCATGCTCTGAA	8072
QY	2021	TCCAGCTACTTGGGAGGGTGAGCGAGGAGAAATTTCTTGAAACCCGGGAGGCGAGGTTGC	2080

Db	8073	TCC	CAGCTATTCAGGAGTTTGAGCGAGAAATCGTTTGAACTTGGAGCGAGAGTTGC	8133			
Qy	2081	AGT	GAAGCGAGATCACACCACTGCACCTCCAGCCTGGGGGA-GAGAGCGAGACTTCCTCTC	2139			
Db	8133	GGT	GAGCGAAGTCGACCAATTCGACTCTAGCTTGGGCAAGAGCGAACTCCGCTCTC	8192			
Qy	2140	AAAA	AAAAACAAAACAAAAGAA	2161			
Db	8193	AAAA	AAAAAAGAGAAAA	8214			
RESULT 5							
ABQ88096/c							
ID	ABQ88096	standard; cDNA; 92638 BP.					
AC	ABQ88096;						
XX	AC						
DT	18-SEP-2002	(first entry)					
XX	DT						
DE	Human osteoblast differentiation related cDNA SEQ ID NO 3.						
XX	DE						
KW	Human; osteoblast; stem cell differentiation; bone tissue deposition;						
KW	osteoporosis; osteopathic; ss.						
OS	Homo sapiens.						
XX	OS						
FN	WO200250301-A2.						
XX	FN						
PD	27-JUN-2002.						
XX	PD						
PF	18-DEC-2001; 2001WO-US048276.						
XX	PF						
XX	18-DEC-2000; 2000US-0255882P.						
PR	24-APR-2001; 2001US-0285691P.						
XX	PR						
PA	(GENE-) GENE LOGIC INC.						
PA	(PROC) PROCTER & GAMBLE CO.						
XX	PA						
PI	Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;						
PI	Mertz L;						
XX	PI						
XX	WPI; 2002-557663/59.						
DR	XX						
XX	XX						
PT	Use of genes and their expression profiles associated with osteoblast						
PT	differentiation for screening modulators bone formation, for diagnosing						
PT	or treating e.g. osteoporosis, or as markers for the differentiation						
PT	process.						
XX	Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.						
PS	XX						
CC	The invention relates to genes and their expression profiles are used						
CC	for: (a) screening modulators of precursor stem cell differentiation into						
CC	osteoblasts, or bone tissue deposition; (b) diagnosing abnormal						
CC	deposition of bone tissue, abnormal rate of osteoblast formation or						
CC	osteoporosis; or (c) treating or monitoring treatment of the conditions						
CC	cited in (b), or monitoring the progression of bone tissue deposition.						
CC	Specific conditions include postmenopausal osteoporosis, glucocorticoid						
CC	osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-						
CC	induced abnormalities in bone formation or bone loss, conditions that						
CC	involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),						
CC	skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome						
CC	or fibrous dysplasia. The present sequence is that of an osteoblast						
CC	differentiation associated cDNA marker of the invention. Note: The						
CC	sequence data for this patent did not form part of the printed						
CC	specification, but was obtained in electronic format directly from WIPO						
CC	at ftp.wipo.int/pub/published_pct_sequences						
XX	XX						
XX	Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;						
Query Match 9.5%; Score 375.4; DB 6; Length 92638;							
Best Local Similarity 55.4%; Pred. No. 5.5e-67;							
Matches 1046; Conservative 0; Mismatches 726; Indels 115; Gaps 12;							

QY 366 ATCAATATTTAGACAGAGTTCTCACTCTGTGCACCAGGCTGGAGTGGACAAAT 425
DB 57931 AGCAATTTTTTTAGACAGAGTTCTCTCTGTGCACTCAGGCTGGAGTGGCCAGT 57872
QY 426 CATAGCTCACTTAACCTCAAACTCCTGGGCTCAAGTGTATCTACTACAGCTCCAG 485
DB 57871 CATAGCTCACTAGAGCTCGACCTTCAGGGCTCAAGCAATCTCTCTGCTCAACCTCCA 57812
QY 486 AGTAGTAGGACTACAGGCACACAGGCCATACCTGGCTAAATTTTTTTTTTAATTTCA 545
DB 57811 AGTAGTGGGACTACAGGCA-GCACCACATATGGCTGGCTAAATTTTACATTT- 57761
QY 546 TTTTATGATATTCATTTTCTTTCTTTTGTGTTGTTGTTGTTGATAGGCTCAGTTT 605
DB 57760 -----TTCTTTTCTTTT-----TTTTTTTTTTTTTTTGGAGAGGAGTCTCGCTCT 57714
QY 606 GTTATCCAGGCTGGAGGAGTGGCATGGTGCAGCTGA--GCAGCCTTTGACTTCTCTGG 663
DB 57713 GTCAACCCAGGCTGAAGTACAGTGTGGATCTTGGCTCACTGCCAGCTCCACCTCCGGG 57654
QY 664 CTCAGTGTATCTCTGCTCAGCTTCCCAAGTAGTGGGACTCAAAACAGTGTCAACA 723
DB 57653 TTCAGGCCATCTCTCTGCTCAGCTTCCAGCTTCCAGTAGTGGGACTACAGGGCCCAACATCA 57594
QY 724 TGCCTGGCTGA-----TATTTTTTTTCTTGAACAGGGTATCACTCTGTTGCCAGGCTG 778
DB 57593 CGCCGAGCTAACTTTTGTATTTTAGTAGACGGGGTTTCAACATGTTAGCCAGGATG 57534
QY 779 GAGTACAGTGGGCTAATAATAGCTCACTGCAGCCTCCCTCC-----TGGGCTCAAGCAATC 835
DB 57533 GTCTCGATCTCTGACCTCGCATCGCCCGCTCGGCTCCCAAGTGTCTGGATTACA 57474
QY 836 CGCTGGCTCAGCATCTGAGTAGTGGGACTACAGGCTTGTGCCACAGGCCAGCTAA 895
DB 57473 GGAGTGAGCCACTATGCTCCGGCTAAAGTTTTTAATTTTTTATAGAGACAAGGTCTCACTA 57414
QY 896 GTTTTAAATAAATGATTTTTGTTATAGAGGAGTCTTGTATGTTGCTCAGGCTGATTTT 955
DB 57413 TTGTTGCCAGGCTGTCTCAAACTTACGTGATCTCTTACCTTAGCTTCCCAAGGGCT 57354
QY 956 TATTTGTGAGACAAGTCTCACTATGTTGCCATGATCCCCACCTCCACTTCCCAAAGT 1015
DB 57353 GGGATACAGGTGGTCTCCCTATTATTAACACTTAGTATGTTGATGTTGTCAAT 57294
QY 1016 GCTCATCTTATCTGTTATGATGAGTGGGACTACAGACATTTAGTGTGTTTCCACTTTTGAC 1075
DB 57293 GAATCATTTAATAACTAACTCCATACTTTTATTTACAGTTTACTCAGTTTCCCTTAATGTC 57234
QY 1076 CATTATGATATATCTCCAGTGAATATTCATGTATACATTTGTGTTGGCATATGTTTCA 1135
DB 57233 CTTTTCCTCAGATCCCAATTTCAATCATCTCTTGGATGATGATGTTTCA 57174
QY 1136 TTTCTGTTGGGTT----- 1149
DB 57173 GACCTTCCCTGTTTTTGTATGACTTACAGTTTTTGGGAGTACCGGTAGGCATATGGTAG 57114
QY 1150 -----ATATCTAGAGTGGAAATGCTGGATCCCGGGTAATATTTTGCAGGAGATT 1202
DB 57113 AATCCCTCTATCAGGATTTGCTGTGATGTTTCTCATGATTTAGACTGAGGTTCTGGGTT 57054
QY 1203 CAGGGGAAGAAACCTTGGGAAATGAAGCATGTTTGAATC--AGCAAGAGTGCAGGG 1260
DB 57053 TTGGAGAGAGAACCAAGAGTCAAGTGGCCATTTTGGCCACATCTTATGAGGGTTCATGC 56994
QY 1261 GTTTTTCGGAGTTTATTTATTTCTGTTGACAAATGTGCAATTTTGTATGAAGATACAG 1320
DB 56993 TGACTTTGATCACTGACTGAGTAACTTTGCCAGACTTCTCCATTTGTAAGTTACTCT 56934
QY 1321 TTATATAAGTAGAAGTGAAGTAATTAAGGCTGAATAGGGCTGTTCAAGATAAATCATGA 1380
DB 56933 TTTTCTTCTTCTGTACTGGACCTTTTGAAGGAAGTCACTGTGTGTCAGGCCACATTTA 56874

QY 1381 AGCACTTTGAATACCAAAATTAAGAGCTTGGCTCTTAACAAAATAATAAAAAATCACAA 1440
DB 56873 AGAATGGGGA-ATTATGTTCTACCTCTTAAGGCGAGAAATATCTACACAGTTTGTGT 56815
QY 1441 TTTTTTTTTTTTTTTTTGAAGAGAGTCTTGTCTTTTCAACCTTGGCTGGAGGCGAGTGTG 1500
DB 56814 TTTTTTTTTTCTTTTGTAGATGGAGTCTCGTCTGTGTTGCCAGGCTGGAGTGCAGTGCA 56755
QY 1501 TGAATCAGCTCAGTCAACTTTTGGCTTCCCGGTTTGAAGCAATTTCTCTGCTTCAGCT 1560
DB 56754 CAATCTGCTCAGTCAAAAGCTCTGCTTCCCGGTTTCAACAATTTCTCTGCTCAGCT 56695
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCAACATGCCAGCTGATTTTGTATTTTAG 1620
DB 56694 CCCGATAGCTGGGACTACAGGCGCCCGCCACATGTCAGCTAATTTTGTATTTTAG 56635
QY 1621 TAGAGATGGGATTTCACTTTTGTGGCCAAAGTGTGCTCAAACTTTTTCG----- 1669
DB 56634 TAGAGACGGGTTTTCACCTTGTAGCCAGGATGGTCTCGACTCTCGACTTGTGATCG 56575
QY 1670 ----TGTCAATATTTGTTGTAATCTTATTTCTTTTGTGCTGAGTGGGCCCCAGACCAA 1725
DB 56574 CCCGCTCGA CTTCAAAGTGTCTGGGATTAACAGCGTGTAGCCACCCGCGCAGTTT 56515
QY 1726 AAAAAATAATCTTAGAATCCAAATCAGTGTGTGTTTGTGACCACTGTCTCACTTTGAGAACA 1785
DB 56514 ATTTGGAATTTCTTCTGCACTGGAGATTTGCTTATTTCTCCCTTATTTATTCATTCAT 56455
QY 1786 CAGTGTACCGAGGCTCAGGAGTAGAGTGTATCTCTGCTCGAAAGAGAAATAGAAATGAA 1845
DB 56454 TTATATCAGCATGGACTTACAGGTATTTTATTTTATCTTTGGGATATGATCTCAAACTAC 56395
QY 1846 AATATTTCTCC-----GGGCCAGGCTGTGTGCTCATGCTCTAATCCAGCACTTT 1896
DB 56394 TTAATTTTGTGCCAAGTGTGCCAGGCTGTGTGCTCACCTGTAGTCCAGCACTTT 56335
QY 1897 GGGAGGCCAAGGCTATGGATCACCTGAGGTCAGGAGTTCAAAACCAAGCTTGGCCAAAT 1956
DB 56334 GGGAGGCCAGGCGAGGCGGATCAC--GAGGTCAGGAGATCAAGACCATCTTGGCTAACAC 56277
QY 1957 GGTGAACCCCGTCTCTACTAAAAATACAAAATAGCTAGTGTGTGGGCGATGCT 2016
DB 56276 AGTGAACCTTGTCTCTATTAATAAATACAAAAAATAGCCAGCATGTTGGCGGCGCT 56217
QY 2017 GTATCCAGCACTCTTGGGAGGCTGAGCGAGAGAAATTTCTTGAAACCGGAGGCGAGG 2076
DB 56216 GTAGTCCAGCACTCTTGGGAGGCTGAGCGAGAGAAATGTTGTAACCCCGGAGGCGGAGC 56157
QY 2077 TTGCAGTGAAGCGAGATCACACCACTGCATCCAGCTTGGGAGGAGAGAGCACTTCT 2136
DB 56156 TTGCAGTGAAGCGAGATCGGCCACTGC-CTCCAGCTGGGTGACAGAGCGAGCTCCAT 56098
QY 2137 CTCAAAAAACAATAAACAAGAAAT 2163
DB 56097 CTCAAAAAGTAAAAATAAATAATTT 56071

RESULT 6

AAK69044/c
ID AAK69044 standard; DNA; 5491 BP.

XX AAK69044;

AC AC

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23856.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

PN

.....

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465566/50.
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX Disclosure; SEQ ID NO 2141; 1180pp; English.
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAS41685-AA42192 represent
CC DNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: the sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 9742 BP; 2382 A; 2564 C; 2516 G; 2280 T; 0 U; 0 Other;
SQ
Query Match 9.3%; Score 366.4; DB 4; Length 9742;
Best Local Similarity 56.3%; Pred. No. 2.3e-65;
Matches 1043; Conservative 0; Mismatches 631; Indels 178; Gaps 12;
QY 370 TATTATTAGAGACAGAGTTCTCACTC-TGTCAACCCAGGCTGGAGTGCAGTGGCAATCAT 428
DB 6351 TTTTCTTTGAGATGAGATTTCATCTTGTGTCACCGCTGGAGTGTGTGGCGCTATCTT 6292
QY 429 AGCTCACTGTAACTCAAACTCCCTGGGCTCAAGTGAATCTTACTACTCAGCTCCAGAGT 488
DB 6291 GGCCTCATGCAACTCCGCTCCCGGGTCAAGCAATTTCTCCTGCCTCAGCCTCTCTGAGT 6232
QY 489 AGCTAGAGCTACAGGACACACAGCCATCTCTGGCTAAATTTTTTTTTTAAATTTTCAATT 548

DB 6231 AGCTGGGTTTACAGGCATGTCCACACGCCCGGCTAAATTTTGTATTTTAA----- 6181
QY 549 TATGTATTCAATTTCTTTCTTTTGTGTTGTTGTTTGTGATAGGCTCTCACATTTGTT 608
DB 6180 -----GTAGAGGCAAGGTTTCTCCATGTT 6157
QY 609 ACCCAGGCTGGAGGCGCAGTGGCATGTGTGACAGCTGAGCAGGCTTTCCTGGGCTCAA 668
DB 6156 GGTCAAGGCTGGT-----CTTGAACCTCCGCACTCAG 6126
QY 669 GTGATCTCTGCTCCCTCAGCCTCCCAAGTAGCTGGGACTACAAACACGTTGTCACCATGCT 728
DB 6125 GTGATCCCGCCCTCCGCTTCCCAAGTGTGGGATTACAGGCTGAGCCACACGCTC 6066
QY 729 GGTGATATTTTTTTTTTCTTGAACACAGGCTATCACTCTCTGTGCCAGGCTGGAGTACAGTG 788
DB 6065 GGCCTTTTTTTTTTATTTTTTGAGATGAGAGTTTCGCTCTCTCACCCAGACTGGAGTCAAGTG 6006
QY 789 GGTAAATAATAGCTCACTGCAAGCTTC-CCCTCTGGGCTCAAGCAATCCGCTGGCTCAG 847
DB 6005 GTGCAGTCTCAACTTACTGCAACCTCTGCTCTCTAGGTTCAAGCAATTTCTCATGCTCAG 5946
QY 848 CATCTCTGAGTAGCTGGGACTACAGCTTGTGCCACAGCCAGCTAAAGTTTAAATAAAT 907
DB 5945 CTTCCCAAGTAGATGGGACTACAGCATGTGCCACACGCTGGCTGATTTTATA--- 5890
QY 908 GATTTTTGGTATAGAGAGGCTCTTCTATGTGCTCAGGCTGTATTTTATTTGTTGAGAC 967
DB 5889 TTTTTTTAGTAGAGACAGTGGTTTCGCTCTGTGCTAGGCTGGTGTGAAACGCTGACT 5830
QY 968 AAGGCTCCTATATGTTGCCATGATCCCGCCAGCTCCTCACTTCCCAAGTGTCTATCTATC 1027
DB 5829 CAAG-----TGATCCGCCACCTCCGCTTCCCAAGTGTGGGATTATA 5786
QY 1028 TGTTCATTAGTCAGTTGACAGACATTTAGGTTGTTTCCACTTTTTCACCATTAATATAA 1087
DB 5785 GCGGTGAGCCAGGCACTGAGTGTGAATTAATTTTAAAGGCCAGTCAATGATCTAAT 5726
QY 1088 TACTCCAGTGAATATTCATGTATATACATTTGTGTGGGCATATGTTTTCATTTCTGTTGGT 1147
DB 5725 GTCTCTGGGTTAATTCATGACCGCCCTTGGTTTATGACAGGCTTCTAGCCATAGCC 5666
QY 1148 TTATATCTAGAGTGGAAATGCTGTGATCCCGGTTATATTTGACAGGCGAGTTCAGGG 1207
DB 5665 CATTTTACAGATAGGACCCAGTCCAGTCTTAGGAAAT----- 5629
QY 1208 GAAGAAAACTTGGGAAATGAAGCATGTTTAGAAATCAGCAAG-AGTGCAGGGGTTTTT 1266
DB 5628 -TAGAGNATTTAGGNACTTGTAACTGTCACAGCCAGGAAATATAGCAGTAAAGAA 5570
QY 1267 CGAGGTTTATTTTATTTATTTCTGTTGACAAATGTGCAAGTTTGTATGAAGATCAAGTTATAC 1326
DB 5569 AAGTCTTGCAGCTAGGCACTGGAGCTCACACCTGTAATCTCAGTGTCTTTGGGAGGCTGAG 5510
QY 1327 TAAGTGAGAAGTGAGAAATTAAGCTGGAATAGGGCTTCAGAGTAAATCATGAAGCACT 1386
DB 5509 GCAGGAGGATCGCTTGAGCCAGGAGTTTGAGACCCAGCTGGGAAACATAGGAGACCT 5450
QY 1387 TTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATAAAATATCAAAATTTTTT 1446
DB 5449 ATCTCTATGAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5390
QY 1447 TTTTTTTTTTGAGAAAGAGTCTTGTCTCTTTTCACTCCCTGGCTGGAGGCGAGTGTGATCT 1506
DB 5389 TTTTTTTTTTGAGACAGAAATCTCACTCTGTGCCAGGCTGGAGTGAATGGTGTATCT 5330
QY 1507 CAGCTCACTGCAACTTTTCGCTCCCGGTTTCAAGCAATTTCTCTGCTTCCAGCTCCCAAG 1566
DB 5329 CAGCTCACCAACCTCCCGCTCTCAGGTTCAAGTAATTTCTCTCCTCAGCTCCTGAG 5270
QY 1567 TAGCTGGGACTACAGGCACTTCCACCATGCGCAGCTGATTTTGTATTTTATTTAGTAGAGA 1626
DB 5269 TAGTGGGACTACAGGCGGTTGCCACCATGTCCGGCTAATTTTGTATTTTATTTAGTAGAGA 5210


```
Db 20447 ATCATACGGCGCGCGCGTGGCTCAGCGCTGTAATCCAGCACTTTGGAGCGCGAG 20506
QY 1909 CATGTGGATCACTGAGGTGAGAGTTCAAAACGAGCCCTGGCCACATGGTGAACCCCG 1968
Db 20507 TGGCGGGATCAC--GAGGTGAGGATCGAGACCATCTCTGGCTAAACGCTGAACCCCG 20564
QY 1969 TCTCTACTAAATAACAAAAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCAGCT 2028
Db 20565 TCTCTACTAAATAACAAAAATAGCCGGGGTAGTGGCGGCGCTGCTCCAGCT 20624
QY 2029 ACTTGGGAGGCTGAGGAGGAGATTTCTTGAACCGGAGGAGGAGGTTGCAAGTGAAGC 2088
Db 20625 ACTCGGAGGCTGAGGAGGAGATGCGCTGAACCGGAGGAGGAGCTTGCAGTGAGCC 20684
QY 2089 GAGATCACACCACTGACATCCAGCTGGGGGAGAGGCGAGACTTCTCTCAAAAAACA 2148
Db 20685 AAGATCGCCCACTGCCCTCCAGCTGGGCGAGAGGCGAGACTCGTCTCAAAAAAAA 20744
QY 2149 AAAAAACAAAGAAATTAAGCAAAATTAGACATTCAGAGAGAA 2189
Db 20745 AAAAAACAAAGAAATTAAGCAAAATTAGACATTCAGAGAGAA 20785
```

RESULT 9

AAD14749
ID AAD14749 standard; DNA; 12394 BP.

AC AAD14749;

XX 01-NOV-2001 (first entry)

XX Human glycogen synthase kinase 3 alpha genomic DNA.

XX Human; glycogen synthase kinase 3 alpha; antidiabetic; cytostatic;
XX antisense therapy; diabetes; hyperproliferative disorder; inflammation;
XX neurological disorder; tumour; haematopoietic disorder; infection;
XX hyperproliferative disorder; developmental disorder; ds.
XX Homo sapiens.

XX Key Location/Qualifiers

```
FT CDS 115..11786
FT /tag= a
FT /product= "Human glycogen synthase kinase 3 alpha"
FT exon 115..397
FT /tag= b
FT intron 398..2437
FT /tag= c
FT exon 2438..2625
FT /tag= d
FT intron 2626..5638
FT /tag= e
FT exon 5639..5722
FT /tag= f
FT intron 5723..5863
FT /tag= g
FT exon 5864..5974
FT /tag= h
FT intron 5975..7901
FT /tag= i
FT exon 7902..8032
FT /tag= j
FT intron 8033..8120
FT /tag= k
FT exon 8121..8227
FT /tag= l
FT intron 8228..9196
FT /tag= m
FT exon 9197..9294
FT /tag= n
FT intron 9295..9374
FT /tag= o
```

```
FT exon 9375..9470
FT /tag= p
FT intron 9471..9897
FT /tag= q
FT exon 9898..10084
FT /tag= r
FT intron 10085..10430
FT /tag= s
FT exon 10431..10523
FT /tag= t
FT intron 10524..11712
FT /tag= u
FT exon 11713..11786
FT /tag= v
XX WO200152865-A1.
XX 26-JUL-2001.
XX 16-JAN-2001; 2001WO-US001411.
XX 21-JAN-2000; 2000US-00488856.
XX (ISIS-) ISIS PHARM INC.
XX Monia BP, McKay R, Butler MM, Wyatt JR;
XX WPI; 2001-442247/47.
XX P-PSDB; AAE06521.
XX Antisense compound 8 to 30 nucleobases in length comprising a compound
XX that is targeted to a nucleic acid molecule encoding glycogen synthase
XX kinase 3 alpha, useful for the treatment of e.g. diabetes and
XX hyperproliferative disorders.
XX Example 15; Page 92-100; 115pp; English.
XX The invention relates to an antisense compound 8 to 30 nucleobases in
XX length targeted to a nucleic acid encoding glycogen synthase kinase 3
XX alpha. The antisense compound specifically hybridises with and inhibits
XX the expression of glycogen synthase kinase 3 alpha. The antisense
XX compound is useful for the treatment of a diseases associated with
XX glycogen synthase kinase 3 alpha such as diabetes, a neurological
XX disorder, a haematopoietic disorder, a hyperproliferative disorder or a
XX developmental disorder. The antisense compounds may also be used
XX prophylactically to prevent or delay infection, inflammation or tumour
XX formation. The present genomic DNA sequence encodes human glycogen
XX synthase kinase 3 alpha protein
XX Sequence 12394 BP; 2780 A; 3043 C; 3547 G; 3024 T; 0 U; 0 Other;
XX Query Match 8.8%; Score 349.8; DB 4; Length 12394;
XX Best Local Similarity 55.2%; Pred. No. 6.4e-62;
XX Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;
QY 370 TATTATTAGACAGAGGTTCTCACTCTGTCACCAGCTGGAGTCAGTGGCACATCATTA 429
Db 3604 TTTTITTTTGAGATGGAGTCTTACTCTGTCACCAGGCTGGAGTCAGTGGCAGTCTTG 3663
QY 430 GCTCACTGTAACCTCAAACTCTGGGCTCAAGTGAATCTTACTCTCAGCTCCAGAGTA 489
Db 3664 GCCCGCTCAACCTCTGCTCCAGGTTCAAGCGGATTTCTGTGCTCAACCACTTGAGTA 3723
QY 490 GCTAGGACTACAGGCACACAGCCATACCTGGCTAAATTTTTTTTTTAAATTTTCATTTT 549
Db 3724 GCTGGATTGAGGCTATGGCCATATGCTGGCTAAATTTTTTTTGTGTAATT--TTTA 3781
QY 550 ATGTATTCAATTTCTTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Db 3782 GTAGAGACAGGGTTTTTGGCATGTTGGCCAGGCTGTCTCGAACTCCTGACCTCAAGTAAT 3841
QY 610 CCCAGGCTGGAGGCGAGTGGCATGGTGACACTGAGCAGCCTTGACTTCTCTGGCTCAAG 669
```

Db 3842 CCGCCCCACCTCGACCTCCCAAAAGTCTGGTATTACAGGCATGGA-----GCCAGCAT 3894
Qy 670 TGATTCCTCTCGCCTCAGCCTCCCAAGTAGCTGGGACTACAAAACACGTGTCA-CCATGCCT 728
Db 3895 ACCTAGCCTGATTTTCAAGACAACTGAAACCTGGATTAGATGTGAATCTTTTTT 3954
Qy 729 GGCTGATATTTTTTTCTTGAACACAGGATATCACTGTGTCGCCAGGCTGGAGTACAGTG 788
Db 3955 TTTTTTTTTTTTTTTTGGAGACGGAGTCTCATGTCTGTCACCCAGGCTGGAGTGTGGTG 4014
Qy 789 GCGTAATAATAGTCACTCGAGCCTCC-CTCCTGGGCTCAAGCAATCCGCTGSCCTCAG 847
Db 4015 GCGTGAATCTCGGTCATCGCAACTCCGCTGCGCGGTTCAGGGATTCCTGTCCTCAG 4074
Qy 848 CATCCTGAGTAGCTGGGACTACAGGCTTGTGCCACAGGCCACGCTAAGTTTTTAAAAAAT 907
Db 4075 CTTCCCTAGTAGCTGGGACTACAGCGGTGTGCCACCACTCTCGGCTAATTTTTTGTAAT 4134
Qy 908 GATTTTGGTATAGAGAGGCTCTGTGATGTTGTCTCAGGCTGTATTTTATTTGTTG--- 963
Db 4135 CTAGTAGGACGGGGTTTACCGAGTTAGCCAGGATGGTCTCTATTTTTTTTTTTTTTT 4194
Qy 964 ---AGACAGGCTCTACTATGTTGGCATGATCCCCCACTCCACTTCCCAAGTCTC 1019
Db 4195 TTTAAGACAGAACTCTGTTCTGTCACTAAGCTGGAGTGCAGTGGTGTGATGTCGGCTCA 4254
Qy 1020 ATCTTATCTGTTCATTAGTCAGTTGACAGACATTTAGGTGTTTCCACTTTTTCACCAAT 1079
Db 4255 CTGCAACTCTGCTCTCTGGTTCAGCGCTGCAACCTCTGCTCTCTGGTTCAGCAAT 4314
Qy 1080 ATGAATAATACCTCAGTGAATATTATGATATACATTTGTGTGGGCAATGTTTTTCATTT 1139
Db 4315 TCTTGTACCTCATCCACTGAGTAGTTGGAATCACAGCGTGCAGCCCACTGCCCCAGCTA 4374
Qy 1140 TGTGGGTTTATATCTAGGAGTGAATTTGCTGGATCCGGGTATATTTTACAGGACGA 1199
Db 4375 ATTTTTTTTGTATTTTATAGTAGAGTGGGTTTTCACACGTTGGCCAGGCTGGTCTCGAAC 4434
Qy 1200 GTTCAGGAGGAAGAAA-----ACTTGGGAAAAATGAAGCATGTTTGAATAATCA 1246
Db 4435 TCCTGGCTCAGTGATTTCCCTGCCTGGCTCCCAAGTGCTGGGATTACAGGGGTGA 4494
Qy 1247 GCAAGATGACAGGGGTTTTTCGGAGTTTTATTTATATTCGTTTGACAAGATGTCAGTTT 1306
Db 4495 GCCACCATGCCAGCTGTTTTTATTTTTTATTTTTTAAAG--GCTGGGTATGTTAGCT 4552
Qy 1307 GATGAAGTACAGTTATATCTAAGTGAAGTGAAGTAAAGGCTGGAATAGGCGCTTCA 1366
Db 4553 CATGCCGTATCTTGAACCTTGAGAGCC--CGAGGACAGGAGGATGCTGAGACTAGGA 4611
Qy 1367 GAGTAAATCATGAAGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATA 1426
Db 4612 GTTCAAAACCACTGGCCACATAGCAGGTTCTTTTAAATAATAATAATAATAAT 4671
Qy 1427 ATAAAAATCAAAATTTTTTTTTTTTTTTTGGAGAAAGATCTGCTCTTTCACCTGGCT 1486
Db 4672 TTTATCTTATTTATTTATTTATTTATTTTGGAGACAGAGTCTGTGCGCCAGGCT 4731
Qy 1487 GGAGGCGAGTGTGTATCTCAGCTCACTGCAACTTTTCGCTCCCGGTTCAAGCAATTC 1546
Db 4732 GGAGTGAGTGGCGGATCTCAGTCTCACTGCAAGCTCCGCTCTCTGGGTTCAAGCCATTC 4791
Qy 1547 TCCTGCTTCAAGCTCCCAAGTAGCTGGGACTACAGGCACTTTCCACCATGCCCCAGCTGA- 1605
Db 4792 TCCTGCTTCAAGCTCCCGAGTAGCTGGGACTACAGGTGCTGCCACCACTGCTGGCTAAT 4851
Qy 1606 ---TTTTGTATTTTATAGATGGGATTTCACTTTGTTGGCCAAAGCTGTGCTCAAACT 1663
Db 4852 TTTTTTTTGTATTTTATAGTACAGGCTTTTCAACCGTGTAAACAGGATGCTCTCAATCT 4911
Qy 1664 TTTTGTGTCTAATTTGTTGTAATTTGTTTCTTTTGTGAGTGGGCCCCAGACCA 1723
Db 4912 CCTGAC-----TTCGTGATCCACCACTCAGCCTCCCAAAAGTCTGGGATTACAGGCGT 4966

Qy 1724 AAAAAATTAATCTTAGAATCCAAATCAATCAGTGTGTGTGTTTGACCACTGTCACTTGAGAAC 1783
Db 4967 GAGCCACACGCTGGCCCT-----GGCTATCTCTTTTAAACCTTTATTTTGGAGAA 5019
Qy 1784 CACAGTGTGACAGGGCTCAGGAGTAGAGTGTCTCTGCTGAAAGAGAAATAGATG 1843
Db 5020 AAAAAATCAGAAAGTGGCATTGTCCTTTTACATGTCAGCAATAAGTTGAAAAAATTTTT 5079
Qy 1844 AAAATATTTCTCCGGCCAGGCTGGTGGCTCATGCTGTAAATCCAGCACTTTTGGGAGGC 1903
Db 5080 TTTTAACTGGGTGGCTGGGCGGTGCTCAGCCCTGTAAATCCAGCACTTTTGGGAGGC 5139
Qy 1904 CAAGGCAATGTGATCA-----CCTGAGGTGAGAGT 1934
Db 5140 TGAGGCTGTGGATCATGAGGTGAGGAGGCTGAGGAGGCTGAGATCAAGAGTCAGGAGA 5199
Qy 1935 TCAAAACCACTGGGCAACATGCTGAACCCGCTCTCTACTTAAATAACAAAAATTAG 1994
Db 5200 TCAGAGCACTCTGCTGCTAAGTGTGTGAACCCCATCTCTACTTAAATAAC-AAAAATAG 5258
Qy 1995 CTAAGTGTGTGGCGCATGCTGTAAATCCAGGCTACTTTGGAGGCTGAGGAGGAGAT 2054
Db 5259 CTGGGCTGTGTGTGCTGCTGTAAATCCAGCTACTTTGGAAGCTGAGGAGGAGAT 5318
Qy 2055 TCTTGAACCCGGAGGAGGAGGTTGCAAGTGAAGCGAGATCACACCACTGCACTCCAGCCT 2114
Db 5319 GCTTGACCCAGGAGGAGGAGGTTGCAAGTGAAGCGGATTTGAGCCACTGCAATGCCAGCCT 5378
Qy 2115 GGGGAGAGAGGAGGAGGTTCTCTCAAAAAACAAAAA 2153
Db 5379 -GGCAACAGAGCAAGACTCTGTCTCAAAAAA 5416

RESULT 10
AAL05122
ID AAL05122 standard; DNA; 13216 BP.
XX
AC AAL05122;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7810.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX
OS cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN W020015320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

1501 TGATCTCAGTCTCACTCAACTTTTCGCTCCCGGGTTCAAGCAATTCCTCTGCTTCAGCCT 1560
 3685 CGATCTCAGTCTCACTCAAGAACTCCACTCCCGGGTTCAAGAGATTCCTCTGCTTCAGCCT 3744
 1561 CCACAGTAGCTGGAGTACAGGCACTTCCACCATCCAGCTGATTTTGTGATTTTAA 1620
 3745 CTTGAGTAGCTGGGATACAGGGGTGCACTTACCGTGCCTGGCTAAATTTATATTTTAA 3804
 1621 TAGAGATGGATTTCACTTTGTTGGCAGAGCTGGTCTCAACTTTTGTGTCATATTCG 1680
 3805 TACAGATGGGTTTCCACTTTGTTGGTCAAGCTGGTCTCGAACTCTCATGATTCGG 3864
 1681 TTGTAATATTGTTCTCTTTGCTGAGTAGGGCCCGCAGACCAAAAAAATAATATCTTAG 1740
 3865 TCCCCCTCGACTCCCAAGAGTTGAGATTACAGGCATGAGCCACTGTGCCCAAGCC --- 3920
 1741 AATCCAAATCAGTGTGTGTTGACCACTGTCACTTGAGAACCAAGTGTGACAGGGC 1800
 3921 ---GCTAATTTATTTTGTACAGACAGACTTCTCCATGTTGCCAGGCTGATCTGGAA 3977
 1801 CTCAGAGTAGAGGTATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGGGCC 1860
 3978 CTCCTGGGCGCAAGTGAACCACTTGCCTCGGC---CTCCAAAGTGTGGGATTACAGGCC 4034
 1861 AGCGGTGGTGGCTCATGCTGTATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCAC 1920
 4035 GGGTGGTGGCTCACGCCCTGTATTCAGACATTTGGGAGGCCAAGGCATGTGGATCAC 4094
 1921 CTGAGTTCAGAGTTCAAAAACAGCCCTGGCCAAATGTTGAAACCCCGTCTCTACTAAAA 1980
 4095 CTGAGTTCAGAGTTCAGACCCAGCCCTGGCCAAAGTGTGACACCCCTGTCTACTAAAA 4154
 1981 ATACAAAAATTAGTAAAGTGTGGTGGCGCATGCTGTATCCAGTACTTGGGAGGT 2040
 4155 ATAC-AAAAATTAGCCACGCTGTGGTGGCGCATGCTGTATCCAGTACTTGGAGGCT 4213
 2041 GAGCGAGGAGAAATTTCTTGAACCCGGGAGGCGAGGTTGAGTGAAGCGAGATCACACCA 2100
 4214 GAGCGAGGAGAAATCACTTGAACCCAGGAGGCGAGGTTGAGTGAAGCGAGATCACACCA 4273
 2101 CTGCACTCCAGCCTGGGGAGAGAGCGAGACTTCTCTCAAAAAAACAACAAA 2160
 4274 CTGCACTCCAGCCTGGGGAGAGAGTGGGCTGTCTCAAAAAACAACAAA 4333
 2161 ATTAAGCAAAATTAGACA 2177
 4334 ACAAACAACAACAACA 4350

RESULT 11

ABL98014
 ID ABL98014 standard; DNA; 13216 BP.

XX AC ABL98014;

XX AC ABL98014;

XX 21-JUN-2002 (first entry)

DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2666.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disease; infection; cytostatic; gene; ds.

XX Homo sapiens.

XX WO200155317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001329.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226271P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228242P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0232402P.
 PR 14-SEP-2000; 2000US-0232403P.
 PR 14-SEP-2000; 2000US-0232404P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.

RESULT 12
ABV52505/C
ID ID ABV52505 standard; cDNA; 396 BP.
XX
AC ABV52505;

DK
XX
XX
PT
PT
XX
PS

XX DT 17-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 52496.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA Schlegel R, Endege WO, Monahan JE;
XX WIPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 10183; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 396 BP; 125 A; 96 C; 57 G; 118 T; 0 U; 0 Other;

Query Match 8.8%; Score 348.6; DB 5; Length 396;
Best Local Similarity 98.9%; Fred. No. 5.1e-62;
Matches 351; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1184 TATTTTGACAGGAGTTCAGGGGAGAAAACCTTGGGAAAATGAAGCATGTTTAGAAA 1243
Db TTTTTCAGGAGAGTTCAGGGGAGAAAACCTTGGGAAAATGAAGCATGTTTAGAAA 329

Qy 1244 TCAGCAAGAGTCAGGGGTTTTCGGAGTTTATTTATATCTGTTGACAAAATGTCAG 1303
Db TCAGCAAGAGTCAGGGGTTTTCGGAGTTTATTTATATCTGTTGACAAAATGTCAG 269

Qy 1304 TTTGATGAAGATACAGTTATCTAAGTGAGAAGTGAAGTAAAGGCTGGAATAGGGCGT 1363
Db TTTGATGAAGATACAGTTATCTAAGTGAGAAGTGAAGTAAAGGCTGGAATAGGGCGT 209

Qy 1364 TCAGAGTAAATCATGAAGCACTTTGAATACAAAATTAAGGAGCTTGGCTGTAAACAAA 1423
Db TCAGAGTAAATCATGAAGCACTTTGAATACAAAATTAAGGAGCTTGGCTGTAAACAAA 149

Qy 1424 ATAATAAAAAATCAAAATTTTTTTTTTTTTTTTTTTTGTAGAAAAGTCTTGTCCCTG 1483
Db ATAATAAAAAATCAAAATTTTTTTTTTTTTTTTTTTTGTAGAAAAGTCTTGTCCCTG 89

Qy 1484 GCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTGCGCTCCCGGTTCA 1538
Db 88 GCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTGCGCTCCCGGTTAA 34

RESULT 13
AAL05121
ID AAL05121 standard; DNA; 29163 BP.
XX AAL05121;
XX 21-NOV-2001 (first entry)
XX Human reproductive system related antigen DNA SEQ ID NO: 7809.
DE Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX Homo sapiens.
XX WO200155320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001339.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226688P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.

PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	DR	WPI; 2001-465570/50.	
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	PT	Isolated nucleic acid molecule encoding a reproductive system antigen is	
PR	29-SEP-2000;	2000US-0236370P.	PT	used in preventing, treating or ameliorating a medical condition.	
PR	02-OCT-2000;	2000US-0237037P.	XX		
PR	02-OCT-2000;	2000US-0237038P.	PS	Disclosure; SEQ ID NO 7809; 1297pp + Sequence Listing; English.	
PR	02-OCT-2000;	2000US-0237039P.	XX		
PR	02-OCT-2000;	2000US-0237040P.	CC	The present invention provides the protein and coding sequences of a	
PR	13-OCT-2000;	2000US-0239335P.	CC	number of human reproductive system related antigens. These can be used	
PR	13-OCT-2000;	2000US-0239337P.	CC	in the prevention and treatment of reproductive system disorders,	
PR	20-OCT-2000;	2000US-0240960P.	CC	including cancer. The present sequence is a genomic sequence encoding a	
PR	20-OCT-2000;	2000US-0241221P.	CC	protein of the invention	
PR	20-OCT-2000;	2000US-0241785P.	XX		
PR	20-OCT-2000;	2000US-0241786P.	SQ	Sequence 29163 BP; 6801 A; 7298 C; 7317 G; 7747 T; 0 U; 0 Other;	
PR	20-OCT-2000;	2000US-0241787P.			
PR	20-OCT-2000;	2000US-0241808P.			
PR	20-OCT-2000;	2000US-0241809P.			
PR	20-OCT-2000;	2000US-0241826P.			
PR	01-NOV-2000;	2000US-0244617P.			
PR	08-NOV-2000;	2000US-0246474P.			
PR	08-NOV-2000;	2000US-0246475P.			
PR	08-NOV-2000;	2000US-0246476P.			
PR	08-NOV-2000;	2000US-0246477P.			
PR	08-NOV-2000;	2000US-0246478P.			
PR	08-NOV-2000;	2000US-0246523P.			
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
PR	17-NOV-2000;	2000US-0249207P.			
PR	17-NOV-2000;	2000US-0249208P.			
PR	17-NOV-2000;	2000US-0249209P.			
PR	17-NOV-2000;	2000US-0249210P.			
PR	17-NOV-2000;	2000US-0249211P.			
PR	17-NOV-2000;	2000US-0249212P.			
PR	17-NOV-2000;	2000US-0249213P.			
PR	17-NOV-2000;	2000US-0249214P.			
PR	17-NOV-2000;	2000US-0249215P.			
PR	17-NOV-2000;	2000US-0249216P.			
PR	17-NOV-2000;	2000US-0249217P.			
PR	17-NOV-2000;	2000US-0249218P.			
PR	17-NOV-2000;	2000US-0249244P.			
PR	17-NOV-2000;	2000US-0249245P.			
PR	17-NOV-2000;	2000US-0249264P.			
PR	17-NOV-2000;	2000US-0249265P.			
PR	17-NOV-2000;	2000US-0249266P.			
PR	17-NOV-2000;	2000US-0249267P.			
PR	17-NOV-2000;	2000US-0249268P.			
PR	17-NOV-2000;	2000US-0249269P.			
PR	17-NOV-2000;	2000US-0249270P.			
PR	17-NOV-2000;	2000US-0249271P.			
PR	17-NOV-2000;	2000US-0249272P.			
PR	17-NOV-2000;	2000US-0249273P.			
PR	17-NOV-2000;	2000US-0249274P.			
PR	17-NOV-2000;	2000US-0249275P.			
PR	17-NOV-2000;	2000US-0249276P.			
PR	17-NOV-2000;	2000US-0249277P.			
PR	17-NOV-2000;	2000US-0249278P.			
PR	17-NOV-2000;	2000US-0249279P.			
PR	17-NOV-2000;	2000US-0249280P.			
PR	17-NOV-2000;	2000US-0249281P.			
PR	17-NOV-2000;	2000US-0249282P.			
PR	17-NOV-2000;	2000US-0249283P.			
PR	17-NOV-2000;	2000US-0249284P.			
PR	17-NOV-2000;	2000US-0249285P.			
PR	17-NOV-2000;	2000US-0249286P.			
PR	17-NOV-2000;	2000US-0249287P.			
PR	17-NOV-2000;	2000US-0249288P.			
PR	17-NOV-2000;	2000US-0249289P.			
PR	17-NOV-2000;	2000US-0249290P.			
PR	17-NOV-2000;	2000US-0249291P.			
PR	17-NOV-2000;	2000US-0249292P.			
PR	17-NOV-2000;	2000US-0249293P.			
PR	17-NOV-2000;	2000US-0249294P.			
PR	17-NOV-2000;	2000US-0249295P.			
PR	17-NOV-2000;	2000US-0249296P.			
PR	17-NOV-2000;	2000US-0249297P.			
PR	17-NOV-2000;	2000US-0249298P.			
PR	17-NOV-2000;	2000US-0249299P.			
PR	17-NOV-2000;	2000US-0249300P.			
PR	17-NOV-2000;	2000US-0249301P.			
PR	17-NOV-2000;	2000US-0249302P.			
PR	17-NOV-2000;	2000US-0249303P.			
PR	17-NOV-2000;	2000US-0249304P.			
PR	17-NOV-2000;	2000US-0249305P.			
PR	17-NOV-2000;	2000US-0249306P.			
PR	17-NOV-2000;	2000US-0249307P.			
PR	17-NOV-2000;	2000US-0249308P.			
PR	17-NOV-2000;	2000US-0249309P.			
PR	17-NOV-2000;	2000US-0249310P.			
PR	17-NOV-2000;	2000US-0249311P.			
PR	17-NOV-2000;	2000US-0249312P.			
PR	17-NOV-2000;	2000US-0249313P.			
PR	17-NOV-2000;	2000US-0249314P.			
PR	17-NOV-2000;	2000US-0249315P.			
PR	17-NOV-2000;	2000US-0249316P.			
PR	17-NOV-2000;	2000US-0249317P.			
PR	17-NOV-2000;	2000US-0249318P.			
PR	17-NOV-2000;	2000US-0249319P.			
PR	17-NOV-2000;	2000US-0249320P.			
PR	17-NOV-2000;	2000US-0249321P.			
PR	17-NOV-2000;	2000US-0249322P.			
PR	17-NOV-2000;	2000US-0249323P.			
PR	17-NOV-2000;	2000US-0249324P.			
PR	17-NOV-2000;	2000US-0249325P.			
PR	17-NOV-2000;	2000US-0249326P.			
PR	17-NOV-2000;	2000US-0249327P.			
PR	17-NOV-2000;	2000US-0249328P.			
PR	17-NOV-2000;	2000US-0249329P.			
PR	17-NOV-2000;	2000US-0249330P.			
PR	17-NOV-2000;	2000US-0249331P.			
PR	17-NOV-2000;	2000US-0249332P.			
PR	17-NOV-2000;	2000US-0249333P.			
PR	17-NOV-2000;	2000US-0249334P.			
PR	17-NOV-2000;	2000US-0249335P.			
PR	17-NOV-2000;	2000US-0249336P.			
PR	17-NOV-2000;	2000US-0249337P.			
PR	17-NOV-2000;	2000US-0249338P.			
PR	17-NOV-2000;	2000US-0249339P.			
PR	17-NOV-2000;	2000US-0249340P.			
PR	17-NOV-2000;	2000US-0249341P.			
PR	17-NOV-2000;	2000US-0249342P.			
PR	17-NOV-2000;	2000US-0249343P.			
PR	17-NOV-2000;	2000US-0249344P.			
PR	17-NOV-2000;	2000US-0249345P.			
PR	17-NOV-2000;	2000US-0249346P.			
PR	17-NOV-2000;	2000US-0249347P.			
PR	17-NOV-2000;	2000US-0249348P.			
PR	17-NOV-2000;	2000US-0249349P.			
PR	17-NOV-2000;	2000US-0249350P.			
PR	17-NOV-2000;	2000US-0249351P.			
PR	17-NOV-2000;	2000US-0249352P.			
PR	17-NOV-2000;	2000US-0249353P.			
PR	17-NOV-2000;	2000US-0249354P.			
PR	17-NOV-2000;	2000US-0249355P.			
PR	17-NOV-2000;	2000US-0249356P.			
PR	17-NOV-2000;	2000US-0249357P.			
PR	17-NOV-2000;	2000US-0249358P.			
PR	17-NOV-2000;	2000US-0249359P.			
PR	17-NOV-2000;	2000US-0249360P.			
PR	17-NOV-2000;	2000US-0249361P.			
PR	17-NOV-2000;	2000US-0249362P.			
PR	17-NOV-2000;	2000US-0249363P.			
PR	17-NOV-2000;	2000US-0249364P.			
PR	17-NOV-2000;	2000US-0249365P.			
PR	17-NOV-2000;	2000US-0249366P.			
PR	17-NOV-2000;	2000US-0249367P.			
PR	17-NOV-2000;	2000US-0249368P.			
PR	17-NOV-2000;	2000US-0249369P.			
PR	17-NOV-2000;	2000US-0249370P.			
PR	17-NOV-2000;	2000US-0249371P.			
PR	17-NOV-2000;	2000US-0249372P.			
PR	17-NOV-2000;	2000US-0249373P.			
PR	17-NOV-2000;	2000US-0249374P.			
PR	17-NOV-2000;	2000US-0249375P.			
PR	17-NOV-2000;	2000US-0249376P.			
PR	17-NOV-2000;	2000US-0249377P.			
PR	17-NOV-2000;	2000US-0249378P.			
PR	17-NOV-2000;	2000US-0249379P.			
PR	17-NOV-2000;	2000US-0249380P.			
PR	17-NOV-2000;	2000US-0249381P.			
PR	17-NOV-2000;	2000US-0249382P.			
PR	17-NOV-2000;	2000US-0249383P.			
PR	17-NOV-2000;	2000US-0249384P.			
PR	17-NOV-2000;	2000US-0249385P.			
PR	17-NOV-2000;	2000US-0249386P.			
PR	17-NOV-2000;	2000US-0249387P.			
PR	17-NOV-2000;	2000US-0249388P.			
PR	17-NOV-2000;	2000US-0249389P.			
PR	17-NOV-2000;	2000US-0249390P.			
PR	17-NOV-2000;	2000US-0249391P.			
PR	17-NOV-2000;	2000US-0249392P.			
PR	17-NOV-2000;	2000US-0249393P.			
PR	17-NOV-2000;	2000US-0249394P.			
PR	17-NOV-2000;	2000US-0249			

08-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249219P.
 17-NOV-2000; 2000US-0249297P.
 17-NOV-2000; 2000US-0249299P.
 17-NOV-2000; 2000US-0249300P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 05-DEC-2000; 2000US-0256719P.
 06-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483232/52.
 Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.
 Disclosure; SEQ ID NO 2665; 766pp; English.
 The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system, reproductive system, immune, respiratory, neurological and cancer, gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a DNA encoding a protein fragment of the invention
 Sequence 29163 BP; 6801 A; 7298 C; 7317 G; 7747 T; 0 U; 0 Other;
 Query Match 8.8%; Score 347.4; DB 4; Length 29163;
 Best Local Similarity 69.9%; Pred. No. 2.4e-51; Indels 11; Gaps 3;
 Matches 515; Conservative 0; Mismatches 211;
 1441 TTTTCTTTTGTGAGAAAGAGCTTGTCTTTTCCACCTGGCTGGAGGGCAGTGGTG 1500
 3623 TCTTTCTTTTGTGAGCTTGTCTTTGTCGCCAGGCTGGAGTGGTG 3682
 1501 TGATCTCAGTCACTGCACTTTCGGCTCCGGGTTCAAGCAATCTCTGCTCAGCT 1560
 3683 CGATCTCAGTCACTGCAAACTCCACTCCCGGGTTCAAGAGATTCTCTGCTTAGCT 3742
 1561 CCCAAGTCTGGGACTAGAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTAG 1620
 3743 CTTGAGTAGCTGGGATTACAGGGCTGCATACCTGCTGGCTGATTTTATATTTTAA 3802
 1621 TAGAGATGGGATTTCACTTTTGTGGCCAGCTGGTCTCAAACTTTTGTCTGTCATAATG 1680

Db 3803 TACAGATGGGTTTCCACCATGTTGTCAGGCTGGTCTGAACTCTGACTTCATGATCCG 3862
 Qy 1681 TTGTAACATATTGTTCTTTTCTGAGGTAGGCCCCCAGACCAAAAAAATAATCTTAG 1740
 Db 3863 TCCCTCTCGACTCCCAAAGTGTAGGATTACAGCATGAGCCACTGTGCCAGCC---- 3918
 Qy 1741 AATCCAAATCAGTGTGTTGGTTTGACCACTCTCACTTGAGAAACCACTGTGACCGAGGC 1800
 Db 3919 ---GCTAATTTATTTTGTACAGACAGATTTCTCCATGTTGCCAGGCTGATCTGNA 3975
 Qy 1801 CTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTCTCCGGGCC 1860
 Db 3976 CTCTGGGCGCAAGTGAACCATCTGCTCGGC---CTCCAAAGTCTGGGATTACAGGC 4032
 Qy 1861 AGCGTGTGCTCATGCTGCTTAATCCAGACACTTTTGGAGGCCAAGCATGTGATCAC 1920
 Db 4033 GGGTCCGCTGCTCAAGCTCTGTAATCCAGCACTTTTGGAGGCCAAGCATGTGATCAC 4092
 Qy 1921 CTGAGGTGAGGAGTTCAAAACAGCCTGGCCAAACATGTTGAAACCCCTCTCTACTAAA 1980
 Db 4093 CTGAGTCCGGAGTTCAAGACAGCCTGGCCAAACAGCTGACACCTCTCTACTAAA 4152
 Qy 1981 ATACAAAAATTAGTAAAGTGTGTCGCGATGCTGTAATCCAGCTACTTTGGAGGCT 2040
 Db 4153 ATAC-AAAAATAAGCCACGCTGCTGTCGCGATGCTGTAATGTCCTACTTTGAGGCT 4211
 Qy 2041 GAGCAGGAGAAATTTCTTGAACCCGGAGGCGAGGTTGCAAGTGAAGCATCACCA 2100
 Db 4212 GAGGAGGAGAAATCACTTGAACCCAGGAGGCGAGGTTGCAAGTGAAGCATCACCA 4271
 Qy 2101 CTGCACTCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA 2160
 Db 4272 CTGCACTCCAGCTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAA 4331
 Qy 2161 ATTAAGCAAAATTAGACA 2177
 Db 4332 ACAAAACAAACAAAAACA 4348
 RESULT 15
 AAS42014/c
 ID AAS42014 standard; DNA; 6437 BP.
 XX
 AC AAS42014;
 DT 17-DEC-2001 (first entry)
 DE Genomic sequence #330 encoding novel human enzyme polypeptide.
 XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
 KW anti arthritic; nephrotropic; anticoagulant; ds.
 XX Homo sapiens.
 OS
 XX WO200155301-A2.
 PN
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001239.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214986P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.

Disclosure; SEQ ID NO 2140; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the

CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 CC infectious disorders (e.g. influenza). The polynucleotides of the
 CC invention can also be used in gene therapy. AAS41685-AAS42192 represent
 CC DNA sequences encoding for the novel human enzyme polypeptides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 6437 BP; 1690 A; 1641 C; 1586 G; 1520 T; 0 U; 0 Other;

Query Match 8.8%; Score 347; DB 4; Length 6437;
 Best Local Similarity 56.7%; Pred. No. 2.1e-61;
 Matches 1052; Conservative 0; Mismatches 625; Indels 180; Gaps 15;

QY 365 TATCATATTATTAGACAGGTTCTCAGTC-TGTACCCAGCGGTGAGTGCAGTGCACA 423
 DB 3050 TTTTCTTTCTTTGAGTGGAGTTTCACTCTTTGTTGCCAGGCTGGAGTGTGCGCGCT 2991
 QY 424 ATCATAGCTCACTGTAACCTCAAACTCCTGGGCTCAAGTGATCCTACTACCTCAGCTCC 483
 DB 2990 ATCTTGGCTCATTTGCACTCTCGCCCTCCCGGGTTCAAGCAATCTCTCGCTCAGCTCC 2931
 QY 484 AGAGTAGCTAGGACTACAGGCACACACAGCCATACCTGGCTAAATTTTTTTTAAATTT 543
 DB 2930 TGAGTAGCTGGGTTTACAGCATGTGCCACACCGCCGGCTAAATTTGTATTTTAA 2875
 QY 544 CAATTTATGATTCATTTCTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 603
 DB 2874 -----GTAGAGGAGGCTTTCTCC 2856
 QY 604 TTGTTACCAGGCTGAGGCGAGTGGCATGGTGACAGCTGAGCGCTTGACTTCTCGGG 663
 DB 2855 ATGTTGGTCAAGCTGGT-----CTTGAATCCCGAC 2825
 QY 664 CTCAGTAGTCTCTCTGCTCAGCTCCCAAGTAGCTGGGACTACAAACAGCTGTACCA 723
 DB 2824 CTCAGGTGATCCCGCGCTCGCTTCCCAAGTGTGGGATTACAGGCTGAGCCACAG 2765
 QY 724 TGCCTGGCTGATATTTTCTTTTCTTGAACAAGGTATCACTGTGCTCCAGGCTGAGTA 783
 DB 2764 CGCTCGGCTTTTCTTTTATTTTGTAGATGGAGTTTCGCTCTGTCTACCCAGAGTGT 2705
 QY 784 CAGTGGCGTAATAATAGCTCACTGCAGCTTC-CCCTCCTGGGCTCAAGCAATCCGCTGGC 842
 DB 2704 CAGTGGTCACTCACTTACTGTGCAACCTCTCGTCTAGGTTCAAGCAATCTCATGC 2645
 QY 843 CTCAGCATCTGAGTAGCTGGGACTACAGCTTGTGCCACCGCCAGCTAAGTTTAA 902
 DB 2644 CTCAGCTTCCCAAGTAGATGGGACTACAGGCATGTGCCACCGCTGGCTAATTTTAT 2585
 QY 903 AAATGATTTTGGTATAGAGGAGTCTTGTGCTATGTTGCTCAGGCTGATTTTATGTT 962
 DB 2584 A-----TTTTTTTATAGAGAGCTGTGCTGTGTTGGCTAGGCTGGTCTGGAACGCT 2529
 QY 963 GAGACAGGCTCTCAGTATGTTGCATGATCCCGCCAGCTCCACTTCCCAAGTGTCAATC 1022
 DB 2528 GACCTCAG-----TGATCCGCCACCTCCGCTCCCAAGTGTGGGA 2485
 QY 1023 TTATCTGTTTCATTTAGTCAGTTGACAGACATTTTAGTGTGTTTCCACTTTTTCACATATG 1082
 DB 2484 TTATAGCGTGCACCGCATCTGAGTGATGAACATTTTAAAGGCCAGTCAATGATC 2425
 QY 1083 AATTAATCTCCAGTGAATATATGATATACATTTGTGTGGGATATGTTTCAATTTCTGT 1142
 DB 2424 TAAATGTCTGTGGGTTAATTCATCTACCGCCCTTGGGTTTATGCAGCAGGTTCTAGCCA 2365
 QY 1143 TGGGTTTATATCTAGGAGTGGAAATGCTGGATCCCGGTAATATTTTGCAGGCAGAGTT 1202

DB 2364 TAGCCATTTTACAGATAGACCACTGAGGATCTTAGGGAAT----- 2323
 QY 1203 CAGGGGAAGAAAACTTGGGAAAAATGAAGCATGTTTGAATAATCAGCAAG-AGTGCAGGGG 1261
 DB 2322 -----TAGAGATTTTAGGAACTTGTCTAACGGTTGCACAGCCAGGAAGTATAGCAGCT 2269
 QY 1262 TTTTTCGAGTTTATTTTATATTTCTGTGTGCAATGTGCAGTTTGTATGAAGATCAAGT 1321
 DB 2268 AAGAAAGTCTTGCAGCTAGGATCGGAGCTCAGACCTGTATCTCAGTCTTGGGAGG 2209
 QY 1322 TATATAAGTGAAGTGAAGATTAAGGCTGGAATAGGCGGTTTCAGAGTAAATCATGAA 1381
 DB 2208 CTGAGGCGAGGAGTCCGCTTGAGGCCAGGAGTTTGAGACCAGCTGGGGAACATAGCGAG 2149
 QY 1382 GCATTTGAAATACCAAAATTAAGGAGCTTGGCTGTAAACAAATAATATAAAATCACAT 1441
 DB 2148 ACCCTATCTCTATGAAAAATAACAAATAAATAAATAGTGTAGAAATATCTGCTTTTCTT 2089
 QY 1442 TTTTCTTTTCTTTTGTGAAAGAGTCTTGTCTTTTACCTGGCTGGAGGCGAGTGTGT 1501
 DB 2088 TTTCTTTTCTTTTGTGACAGAAATCTCAGCTGTGTCAGGCTGGAGTGCATGGTGT 2029
 QY 1502 GATCTCAGCTCACTGCAACTTTTCGCTCCCGGGTTTCAAGCAATTTCTCTGCTTCAGCCTC 1561
 DB 2028 AATCTCAGCTCACCAACCTCCGCTCTCAGGTTCAAGTGATTTCTCTCCCTCAGCCTC 1969
 QY 1562 CCAAGTAGTGGGACTACAGGCATCTCCACCATGCCAGCTGATTTTGTATTTTGTAGT 1621
 DB 1968 CTGAGTAGTGGGACTACAGCGGGTGCACCATGTCCGGCTAAATTTTGTATTTTGTAGT 1909
 QY 1622 AGAGATGGATTTCACTTTTGTGGCAAGCTGTGTCTCAAACTTTT----- 1666
 DB 1908 AGAGACGGGTTTCACTATGTTGACAGGCTGTCTCGACTCTCTGACCTTGATCTGCTG 1849
 QY 1667 -----TGCTGTCAATAATGTTGTAA-----CTATTG 1692
 DB 1848 CGGCTCGGCTCCCAAGAGTGTGGGATTAAGGTTGAGCCACCTCGCTCGGCTCTT 1789
 QY 1693 TTCTTTTGTGAGTGGGCCCCAGACCAAAAAATAAATCTTAGAATCAAAATCAG 1752
 DB 1788 GGTGTTTCTTTTACAAACAGTGGCCAGCCAGATNTTCTTTATTTCCGCGAAGAACCTTC 1729
 QY 1753 TGTGTTGGTTTG-ACCACTGTCACTTGAGAAACACAGTGTGACAGGCGCTCAGGAGT-- 1809
 DB 1728 GGTGAGATACAGACAGTACTCAAGATCAATTGATATAACCGCTAGTTGTAATTC 1669
 QY 1810 --AGAGGTGATCTCTGCTGAAAGAGAAATAGAAATATTTCTCGGCGCCAGCGGTG 1867
 DB 1668 CAACATTTGGATCTGACCTCCAGGATGCACCTGTTAGAAAGTGGATTGGGCGGCGGTG 1609
 QY 1868 --GTGCTCATGCTGTAATCCAGCACTTGGGAGGCCAAGGCATGTGGATCACTGAG 1925
 DB 1608 CTGTGCTCAGACCTGTAATCTAGCGTTTGGAGGAGAGAGCGAGCTGATTTGCTTGA 1549
 QY 1926 GTGAGGTTTCAAAACAGCTGGCCAAACATGTAAGAACCCCGTCTCTACTAAAAATACA 1985
 DB 1548 GTGAGTGTATTAATAGCTGTGTCAACATGTTGAACCTCATCTCTACTAAAAATACA 1489
 QY 1986 AAAAATAGCTAAGTGTGTGGCGCATGCTGTATCCAGCTACTTGGGAGGTTGAGGC 2045
 DB 1488 AAAAATAGCC--AGGCGTGTGGCAACCACTATAGTCCAGCTACTTGGGGGCTGAGGC 1431
 QY 2046 AGGAGATTTCTTGAACCGGAGGAGGAGGTTGAGTGAAGCAGATCACCACTGCA 2105
 DB 1430 AGGAGAACTGCTGTAGCCCGGAGGTGGAGTTGAGGAGGCCAGATC-TGCCACTGCA 1372
 QY 2106 CTCACGCTGGGAGAGAGCGAGCTTCTCTCAAAAAAACAACAAAAAACAAGAAAT 2162
 DB 1371 CTCTAGCTGAGGACAGAGCAAGACTCCATCTCAAAAAAAGAAAAAAGAAAGT 1315

Search completed: May 12, 2004, 19:47:11

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 284.726 Seconds
(without alignments)
7714.406 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattcatccatttaaat.....gggacctccacgcacgcgg 3958

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/6C_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/6D_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	349.8	8.8	12394	4	US-09-488-856A-10
2	345	8.7	84495	4	US-09-797-906-3
3	322.4	8.1	3805	4	US-09-108-006C-3
4	321	8.1	35060	3	US-08-814-095-7
5	319.8	8.1	392000	4	US-10-027-983-11
6	316.6	8.0	59065	4	US-09-813-817-3
7	316.6	8.0	59065	4	US-09-978-157-3
8	312.4	7.9	174493	4	US-09-804-471A-3
9	312.4	7.9	174493	4	US-10-238-709-3
10	308.2	7.8	162450	4	US-09-345-882-1
11	307.6	7.8	43950	4	US-09-735-934A-3
12	307.6	7.8	43950	4	US-10-060-332-3
13	307.2	7.8	70000	4	US-09-851-896-3
14	304	7.7	39982	4	US-09-820-924-3
15	303	7.7	43950	4	US-09-735-934A-3
16	303	7.7	43950	4	US-10-060-332-3
17	301.8	7.6	246240	2	US-08-724-394A-20
18	301.8	7.6	246240	2	US-08-724-394A-21
19	301.8	7.6	246240	2	US-08-724-394A-22
20	299.8	7.6	39982	4	US-09-820-924-3
21	297.2	7.5	70000	4	US-09-851-896-3
22	296.6	7.5	59065	4	US-09-813-817-3
23	296.6	7.5	59065	4	US-09-978-157-3
24	291.2	7.4	83450	4	US-09-811-469-3
25	289.2	7.3	55298	4	US-09-491-356C-1
26	288	7.3	21234	4	US-09-810-671-3
27	288	7.3	21234	4	US-10-109-854-3

28	287.4	7.3	116592	4	US-09-818-512-3	Sequence 3, Appli
29	286.8	7.2	35060	3	US-08-814-095-7	Sequence 7, Appli
30	285.8	7.2	162450	4	US-09-345-882-1	Sequence 1, Appli
31	283.8	7.2	12394	4	US-09-488-856A-10	Sequence 10, Appli
32	282.8	7.1	3805	4	US-09-108-006C-3	Sequence 3, Appli
33	282.8	7.1	246240	2	US-08-724-394A-20	Sequence 20, Appli
34	282.8	7.1	246240	2	US-08-724-394A-21	Sequence 21, Appli
35	282.8	7.1	246240	2	US-08-724-394A-22	Sequence 22, Appli
36	280	7.1	116592	4	US-09-818-512-3	Sequence 3, Appli
37	275	6.9	83450	4	US-09-811-469-3	Sequence 3, Appli
38	275	6.9	84495	4	US-09-797-906-3	Sequence 3, Appli
39	270.4	6.8	72604	4	US-09-268-392-7	Sequence 7, Appli
40	270.4	6.8	72604	4	US-09-657-474-7	Sequence 7, Appli
41	269.4	6.8	23187	4	US-09-499-522-1	Sequence 1, Appli
42	267.2	6.8	45546	4	US-09-146-053-6	Sequence 6, Appli
43	265.6	6.7	21721	4	US-09-269-939A-41	Sequence 41, Appli
44	264	6.7	44848	4	US-09-435-739-42	Sequence 42, Appli
45	263	6.6	98844	4	US-09-791-211-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-488-856A-10

; Sequence 10, Application US/09488856A

; Patent No. 6316259

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Robert McKay

; APPLICANT: Madeline M. Butler

; APPLICANT: Jacqueline Wyatt

; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP

; FILE REFERENCE: RTS-0115

; CURRENT APPLICATION NUMBER: US/09/488,856A

; CURRENT FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 88

; SEQ ID NO 10

; LENGTH: 12394

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (115)...(397)

; NAME/KEY: CDS

; LOCATION: (2438)...(2625)

; NAME/KEY: CDS

; LOCATION: (5639)...(5722)

; NAME/KEY: CDS

; LOCATION: (5864)...(5974)

; NAME/KEY: CDS

; LOCATION: (7902)...(8032)

; NAME/KEY: CDS

; LOCATION: (8121)...(8227)

; NAME/KEY: CDS

; LOCATION: (9197)...(9294)

; NAME/KEY: CDS

; LOCATION: (9375)...(9470)

; NAME/KEY: CDS

; LOCATION: (9898)...(10084)

; NAME/KEY: CDS

; LOCATION: (10431)...(10523)

; NAME/KEY: CDS

; LOCATION: (11713)...(11786)

; US-09-488-856A-10

Query Match 8.8%; Score 349.8; DB 4; Length 12394;

Best Local Similarity 55.2%; Pred. No. 4.9e-77;

Matches 1016; Conservative 0; Mismatches 742; Indels 81; Gaps 14;

Qy 370 TATTATTAGACAGGTTCTCACTCTGTCCAGGCTGGAGTGCAGTGGCAATCAT 429

Db 3604 TTTTITGAGATGGAGTCTCTGTCTGTCCAGGCTGGAGTGCAGTGGCAATCTTG 3663

Query Match 8.78; Score 345; DB 4; Length 84495;
Best Local Similarity 55.04; Pred. No. 2.1e-75;
Matches 9/9; Conservative 0; Mismatches 715; Indels 85; Gaps 12;
QY 445 AAACCTCTGGGCTCAAGTATCTTACTACCTCAGCC-TCCAGAGTAGCTAGGACTACAGG 503
DB 18461 AACTTCTGGGCTGAAGCCATCCACCTGCCTGGCTTCCAAAGTGGTGGGATTACAAG 18402
QY 504 CACACAGGCGATACCTGGCTAATTTTTTTTTTAAATTTTCAATTTATGATTTCAATTTTC 563
DB 18401 CATTAGTCACACGCGCTGGACTATTATATATATATATATATATATATATATATATTTTT 18342
QY 564 TTTCTTTTGTG-TGTGTTGTTTGTAGATAGGCTCTCACTTGTCTTACCCAGGCTGGAGG 622
DB 18341 TTTTTTTTTTTTTTTTTTTTTTTTGTAGACAGGGTCTCACTGTCTACCCAGGCTGGAGT 18282
QY 623 GCAGTGGCATGGTGACAGCTGA--GCAGCCTTGACTTCTGGGCTCAAGTGATCCTCCTG 680
DB 18281 GCAGTGGCATGATCTCAGCTCAGCTGCAACCTTGAGCTCCTGGGTTCAAGCAATTTCTCCTG 18222
QY 681 CCTCAGCTCCAGTAGTGGGACTACAAACAGCTGTACCATGTCCCTGGCTGATATTTT 740
DB 18221 CCTCAGCTCTTAAAGTAGTGGGATTACAGCACATGTACCGCGCCAGCTAATTTTTTT 18162
QY 741 TTTTCTTGAACAGGGTATCACTCTGTGTCAGGCTGGAGTACAGTGGGCTAATAATAG 800
DB 18161 TTTTCTT-----TTTTTGTTTTTTTTTTTTGTAGAGCGGAGTTTCACTATATAG 18109
QY 801 CTCACCTCAGCTCCTCCTGGGCTCAAGCAATCCGCTGGCCTCAGCATCCTGAGTAGC 860
DB 18108 CCAGCGCATCTCGAATCTCTGACCTCAAGTATCTGCCGCTCAGCTCCCAAGTGT 18049
QY 861 TGGGACTACAGGCTGTGCGACAGGCGCCAGCTAAGTTTAAATAATGATTTTGGTATA 920
DB 18048 TAAGATTACAGGCTAGGACCACTGCGCGCTCTCTATTTTTATAGTTTAAAGACACA 17989
QY 921 GAGGAGCTGTGCTATCTGTCTAGGCTGATTTTTTATTTGTAGAGCAAGTCTCACTAT 980
DB 17988 GA-----GTCCACACTCTGTGCGGCTAAGCAATCTCTACCTCAGCTCCGAGTAGCTACATGCAC 17933
QY 981 GTTGCCATGATCCCGCCACCTCCACTTCCCAAGTGTCTATCTTATCTG--TTCAATTAGT 1038
DB 17932 GCCTTAACTCCTGGGCTAAGCAATCTCTACCTCAGCTCCGAGTAGCTACATGCAC 17873
QY 1039 CAGTTGACAGACATTTAGTGTGTTTCCATTTTGGACATATGAATAACTCCAGTGA 1098
DB 17872 ATGCCACCATACCAGCTAATTTTTTAAATGTTTTTAAAGATGGGGTCTCACTATGTT 17813
QY 1099 ATATTCTATGATACATTTGTGCGGCATATGTTTCTATTTCTGTTGGTTTATATCTAGG 1158
DB 17812 CCAGGCTGGTCTGGAATCTGTGCGCTTAAGCGATCTCCATCTTGGCTTCCAAAGTG 17753
QY 1159 AGTGGAAATGCTGGATCCCGGTAATATTTTGAAGAGAGAGTTTTCAGGCGAGGAGGAAACT 1218
DB 17752 CTGGGATTACAAGTGTGAGCCACCATGTCTGGCCACAAATGTACTTTTCAAGGTCACTT 17693
QY 1219 TGGGAAATGAGCATGTTTGAATATCAGCAAGTGTGAGGGGTTTTTCGAGTTTTATT 1278
DB 17692 GACTGCACTGTGG--AGAATAAATCGAGGAAGTATAATAGTTTAAATAGGCGAGTATTTA 17635
QY 1279 TTATATTTCTGTGACAAATGTCAGTTTGTATGAAGATACAAAGTTATTAAGTACAGAGT 1338
DB 17634 ATGACCATACAAAGCCCATCTCACTTATGAATAAATCATATTTACTGATCATTTCT 17575
QY 1339 GAGAAATTAAGGCTGGAATFAGGCGCTTCAGAGTAAAAATCATAGCACTTTGA-----ATAC 1394
DB 17574 CTATTTTATTGGACATTTAGACTATTTCCAGTCTTTGCTTAAATTAATAAAGCGACAC 17515
QY 1395 CAAATTAAGAGCTTGGCTGTAAACAAATAATAAATAATCAAAATTTTTTTTTTTTT 1454
DB 17514 TGTAAATAACATTTCTGAAACCTAAATCTTTGGCAGCATCTCTAATAATTTTCTTCTTTT 17455
QY 1455 TTGAGAAAGAGTCTTGCTCTTTTCCCTCGGCTGGGCGGAGTGGTGTATCTCAGCTCAC 1514

DB 17454 TTTTGAACAGAGCTCTACCTGTTGCCAGGCTCGAGTGGTGGTGAATCCAGCTCAC 17395
QY 1515 TGAACCTTTGCTCCCGGTTCAAGCAATCTCTGCTTCAAGCTTCCAGCTTCCAGTGTGGG 1574
DB 17394 TGAACCTTCCACCTCTCTGGGTTCAAGTGTATCTCTGCTTCAAGCTTCCAGTGTGGG 17335
QY 1575 ACTACAGGACCTTCCACCATGCTCCAGCTGATTTTGTATTTTGTAGAGATGGGATTT 1634
DB 17334 ATTACAGGTTGGGCGCCACCATGCTGCTGCTAATTTTGTACTTTTACAGAGACAGGTTT 17275
QY 1635 C-ACTTGTGTGGCAAGCTGCTCTCAAACTTTT----- 1666
DB 17274 CGCAATTTTGGCAGGCTGCTCTCAAACTTCTGACCTCTGGGATTACAGCGGTAGCCA 17215
QY 1667 -----TGCTGTCAATATGTTGTAACTATTGTTCTTCTGCTGAGGTAGGCGCCCA 1718
DB 17214 CCAGCGCGCACTTAATTTTCTAAATATAGATTTCTAGAGTGGAAATTAAGTAATAA 17155
QY 1719 GACCAAAAAAATAAATCTTAGAATCCAAATCAGTGTGTTGTTGTTGACCTGCTCACTTG 1778
DB 17154 GTTAAAGGTTATGAATTTTGTGGCTATTTTAAATAATTTTAAAGATATAAGATT 17095
QY 1779 AGAACACAGCTGTGACAGGCGCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATA 1838
DB 17094 TTGATCATTAATAATAAATAATCTACTTTGCAAAATGGAATATCTTTTCACTATCCAGTC 17035
QY 1839 GAAT-----GAAATATTTCTCGGCGCAGGCTGGTGGCTC 1874
DB 17034 AAATTTATCAGCGGAAATGCTTTTAAATAAGATGAATAAATACTAGGCGAGCTGGCTG 16975
QY 1875 ATGCTGTATATCCAGCACCTTTGGAGGCGCAAGCATGTGGATCACTGAGGTGAGGAT 1934
DB 16974 ACCTGTATATCCAGCATCTAGGAGGCTGAGGTGGCGGATTC--AAGTCAAGAT 16917
QY 1935 TCAAAACCGCTGGCCAACTGTTGAAACCCCTCTCTACTAAATAAATAAATAAATAA 1994
DB 16916 TCAGAGCATCTCTGGCTAATCATGTGTAACCCCATCTCTACTAAATAAATAAATAA 16857
QY 1995 CTAGTGTGTGGCGCATGCTGTATCCAGCTACTTGGAGGCTGAGCGAGGAAAT 2054
DB 16856 CCGGCTGTGCTGGCGGCGCTGTAGTCCAGCTACTTCGGAGGCTGAGCGAGGAAAT 16797
QY 2055 TCTTGAACCGGAGGCGAGAGTTCAGTGAAGCGAGATCACACCTGCACTGCACTCCAGCT 2114
DB 16796 TCTTGAACCGGAGATGAGGTTGAGTGGCGGATTCAGTGGCGGATTCAGCTCCAGCT 16737
QY 2115 GGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAA 2153
DB 16736 GGGCAACAGGCGCAGACTCTGATTCAAAAAAGAAAAA 16698

RESULT 3

US-09-108-006C-3/c
; Sequence 3, Application US/09108006C
; Patent No. 6524613

GENERAL INFORMATION:

APPLICANT: Steer, Clifford J.
Kren, Betsy T.
Bandyopadhyay, Paramita
Roy-Chowdhury, Jayanta
TITLE OF INVENTION: Hepatocellular Chimeraplasty
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kimeragen, Inc.
STREET: 300 Pheasant Run
CITY: Newtown
STATE: PA
COUNTRY: USA
ZIP: 18940

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible


```
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 27385..27387
; FEATURE:
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 28129..28131
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance
; OTHER INFORMATION: gene"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
```

```
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32386..32468)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
; US-08-814-095-7
```

Query Match 8.1%; Score 321; DB 3; Length 35060;
Best Local Similarity 68.8%; Pred. No. 1.3e-69;
Matches 511; Conservative 0; Mismatches 205; Indels 27; Gaps 4;

```
QY 1445 TTTTCTTTTTCAGAAAGAGTCTTCTTTTCCCTGCGAGGGGAGTGGTGTGAT 1504
      |||||
Db 9802 TCTTTTCTTTTTCAGAGCGGAGTCTCTATCGCCAGGCTGGAGTGCACAAAT 9861
      |||||
QY 1505 CTCAGCTCACTGCAACTTTTCGCGCTCCCGGGTTTCAAGCAATTCTCTCTTTCAGCTCCCA 1564
```


QY 1801 CTCAGAGTAGAGTGATCTCTGCTCGAAGAGAGAAATAGAAATTAATCTTCGGGGCC 1860
 Db 126735 GTCTTCTGAATGTTTATTTCTTCCCTGTTAATACAGTAAATAATTAAGTTGGCC 126794
 QY 1861 AGCGTGTGGCTCATGCTGTAATCCAGCATTTCAGGAGCCAGGCGATGTGATCAC 1920
 Db 126795 AGGAGTGTGGCACAATGCTGTAATCCAGCATTTCAGGAGCCAGGCGAGCATCAC 126854
 QY 1921 CTGAGTCAAGGAGTTCAAAACACAGCTGGCCAAATGCTGAAACCCGCTCTCTACTAAAA 1980
 Db 126855 CTGAGTCAAGGAGTTCGAGACCCAGCTGGCCAAATGCTGAAACCCGCTCTCTACTAAAA 126914
 QY 1981 ATACA----AAAAATPAGCTAAAGTGTGGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 2036
 Db 126915 AAAAATGTAAAGTTAGCTGGGTGTGGTGTGGCGACCTGTAGTCCAGCTATTCCGGGA 126974
 QY 2037 GGGTGAAGGAGGAGAAATTTCTTAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 2096
 Db 126975 GGGTGAAGGAGGAGAAATTTCTTAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 127034
 QY 2097 ACCACTGCACTCCAGCTGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 2156
 Db 127035 GCCACTGCACTTAGCTGGGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 127093
 QY 2157 AAGAATTAAGCAAAATPAGCATTTGAGAGAGAACTGAAAGGGGTGAGACCGATCACAG 2216
 Db 127094 AGTTAAGTAGTACTTTGGGCGCTATCAGATAGTGTCTCTCTGAGGAGGCTGGGCGCCCTA 127153
 QY 2217 TTCTGTGCCACAT 2230
 Db 127154 CTACAGTTTCACTT 127167

RESULT 9

US-10-238-709-3
 ; Sequence 3, Application US/10238709
 ; Patent No. 6680188
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marlon et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001164DIV
 ; CURRENT APPLICATION NUMBER: US/10/238,709
 ; CURRENT FILING DATE: 2002-09-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 174493
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(174493)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-238-709-3

Query Match 7.9%; Score 312.4; DB 4; Length 174493;
 Best Local Similarity 61.7%; Pred. No. 4.4e-67;
 Matches 490; Conservative 0; Mismatches 298; Indels 6; Gaps 3;

QY 1441 TTTTCTTTTCTTTTGAAGAGTCTTCTTTCACCTGGCTGGAGGCGAGTGGTG 1500
 Db 126376 TCTCTTTCTTTCTTCAAGACAGAGTCTTCTCTGTCCACCGAGCTGGAGTGGAGTG-G 126434
 QY 1501 TGATCTCAGTCTCACTGCAATTTGGCTCCCGGGTCAAGCAATTTCTCTGCTTCAGCCT 1560
 Db 126435 CGACCTTGGCTCACTGCAACCTCGGCTCTCTGGTTCAGAGAGATTCTCTGCTCAGCTC 126494
 QY 1561 CCCAAGTAGCTGGAGCTACAGGAGCTTCCACCATGCCAGCTGATTTTGTATTTTAG 1620
 Db 126495 CCCGAGTAGCTGGAGTTACAGGTGGCGCCGCCACCATGCTGGCTAATTTTGTATTTTAG 126554

QY 1621 TAGAGATGGGATTTTCACTTTGTGTGGCCAAAGCTGGTCTCAAACTTTTGTGTCTAATTTG 1680
 Db 126555 TAGAGATGGGTTTCAACCATTTGAAGCATTTTAGAAAGGGGTTAGGAGGAGAGACA 126614
 QY 1681 TTGTAACTATTCTTCTTCTTGTCTGAGTGGGCCCCCAGACCAAAAAATAAATCTTAG 1740
 Db 126615 NNN 126674
 QY 1741 AATCCAAATCAGTGTGTTTGGTTTGACCACTGTCACTTGAGAACCAACAGTGTGACCAAGGC 1800
 Db 126675 NNTCTCCCCCTTTCCTCCAAAATATGGCAGACTCTTCTCTCCCTAGTCTCATATA 126734
 QY 1801 CTCAGAGTAGAGTGATCTCTGCTCGAAGAGAAATAGAAATTAATTTCTCCGGGCC 1860
 Db 126735 GTCTTCTGAATGTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 126794
 QY 1861 AGCGTGTGGCTCATGCTGTAATCCAGCATTTCAGGAGCCAGGCGATGTGATCAC 1920
 Db 126795 AGGAGTGTGGCAGATGCTCTGTAATCCAGCATTTCAGGAGCCAGGCGAGCATCAC 126854
 QY 1921 CTGAGTCAAGGAGTTCAAAACACAGCTGGCCAAATGCTGAAACCCGCTCTCTACTAAAA 1980
 Db 126855 CTGAGTCAAGGAGTTCGAGACCCAGCTGGCCAAATGCTGAAACCCGCTCTCTACTAAAA 126914
 QY 1981 ATACA----AAAAATPAGCTAAAGTGTGGGCGCATGCTCTGTAATCCAGCTACTTTGGGA 2036
 Db 126915 AAAAATGTAAAGTTAGCTGGGTGTGGTGTGGCGACCTGTAGTCCAGCTATTCCGGGA 126974
 QY 2037 GGGTGAAGGAGGAGAAATTTCTTAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 2096
 Db 126975 GGGTGAAGGAGGAGAAATTTGTAACCCGGGAGGAGAGTTCAGTGAAGCGAGATCAC 127034
 QY 2097 ACCACTGCACTCCAGCTGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAA 2156
 Db 127035 GCCACTGCACTTAGCTGGGAGAGAGTTCCTCTCAAAAAACAAAAACAA 127093
 QY 2157 AAGAATTAAGCAAAATPAGCATTTGAGAGAGAACTGAAAGGGGTGAGACCGATCACAG 2216
 Db 127094 AGTTAAGTAGTACTTTGGGCGCTATCAGATAGTGTCTCTCTGAGGAGGCTGGGCGCCCTA 127153
 QY 2217 TTCTGTGCCACAT 2230
 Db 127154 CTACAGTTTCACTT 127167

RESULT 10

US-09-345-882-1/c
 ; Sequence 1, Application US/09345882
 ; Patent No. 6399373
 ; GENERAL INFORMATION:
 ; APPLICANT: Bouguieret, Lydie
 ; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
 ; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
 ; FILE REFERENCE: GENSET.031A
 ; CURRENT APPLICATION NUMBER: US/09/345,882
 ; CURRENT FILING DATE: 1999-06-30
 ; PRIOR APPLICATION NUMBER: US 60/091,315
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/111,909
 ; PRIOR FILING DATE: 1998-12-10
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 1
 ; LENGTH: 162450
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 72794
 ; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
 ; FEATURE:
 ; NAME/KEY: allele


```
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
```

```

;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103783..103828
; OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
;
; FEATURE:

```

```

Query Match      7.8%; Score 308.2; DB 4; Length 162450;
Best Local Similarity 66.7%; Pred. No. 4.8e-66;
Matches 529; Conservative 0; Mismatches 233; Indels 31; Gaps 5;

Qy 1421 AAAATAATAAAAAATCAAAATTTTTTTTTTTTTTTTGTGAGAAAGAGTCTTGCTCTTTCCACC 1480
Db 86851 AATCTAAACAAAGTGAACAACTATTTTCTTTCTTTTGTAGGGGAGTCTAGCTCTGT-TGC 86793

Qy 1481 CTGGCTGAGGGCAGTGGTGTGATCTGCTCAGCTCAGTGCACACTTGGCTCCCGGGTTCAAG 1540
Db 86792 CGGGCTGGAGTGCAGTGGCCACCACTTGGCTCAGTGCACACTTGGCTCCCGGGTTCAAG 86733

Qy 1541 CAATTTCTCTGCTTTAGCTCCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGGCCA 1600
Db 86732 CAATTTCTCTGCTTTAGCTCCCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGGCCA 86673

Qy 1601 GCTGATTTTGTATTTTAGTAGAGATGGGATTTTCACTTTTGTGGCCAGCTGCTCTCAA 1660
Db 86672 GCTAAATTTTGTATTTTGTAGAGAGAGGGGTTTCAACCATTTAGCCAGGATAGTCTCGA 86613

Qy 1661 ACT-----TTTGTGCTGCATAAATGTTGTAACTATTTCTTTTGTGAGGT 1708
Db 86612 TCTCTGACCTGCTGATCCGGCCCACTTGGCTCCCAAGTGTGGGATTAAGCGGTGAG 86553

Qy 1709 AGGCCCCCAGACACAAAAATAAAATCTTAGAATCCAAATCAGTGTGTTGGTTGACCA 1768
Db 86552 CCACCCCACTGGCCAAAAAGTAAAACTATTATAATAATTAAGCTTTTAAAGTGA 86493

Qy 1769 CTGTCACTTGAGAACCAAGTGTG-----TGACAGGGCCCTCAGGATAGAGGTGATCTC 1821
Db 86492 CATGGGAATACCAACCCCTTGTATTTCAATGATCAGTCAAAATCTTAAATAATTTTCCC 86433

Qy 1822 TGCTCGAAGAGAAATAGAAATGAAATATTTCTCGGGCCAGG-----CGTGTGG 1871
Db 86432 TGGCAGAAACAAAGCAATCTTCAAGCTTCTATTAAGATGATGATGATTTTGCT 86373

Qy 1872 CTCAATGCTGTAAATCCAGCACTTTGGGAGGCCCAAGCATGTGATCACTTGGAGTCAAG 1931
Db 86372 GGAAGTATTGTACTTCCAGCACTTTGGGAGGCCCAAGCATGTGATCACTTGGAGTCAAG 86313

Qy 1932 AGTTCAAAACAGGCTGGCCCAAGTGTGAAACCCCGTCTCTACTAAATAATACAAAAAAT 1991
Db 86312 AGTTTCGAGACCAAGCTGGCCCAAGTGTGAAACCCCGTCTCTACTAAATAATACAAAAAAT 86254

```

```

Qy 1992 TAGCTAAGTGTGGTGGCGCATGCTCTGTAATCCAGCTACTTGGGAGGCTGAGGCGAGGA 2051
Db 86253 TAGCCAGGCATGGTGGCGGCTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGCGAGGA 86194

Qy 2052 ATTTCTTGAACCCCGGAGGCGAGGTTGCACTGGAAGCGAGATCACACCACTGCACTCCAG 2111
Db 86193 ATTGCTGAACCCCGGAGGCGAGGTTGCACTGAGCCGATATCGCACCACTGCACTCCAG 86134

Qy 2112 CTGGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAAAGAAATTAAGCAAT 2171
Db 86133 CCTGGGCGACAGAGTGAAGCTCCGCTCAAAAAACAAAAACAAAAAGATGATTGTACTTCAA 86074

Qy 2172 TAGACATTGCGAGA 2184
Db 86073 TGGCACTTAAAGA 86061

```

RESULT 11

```

US-09-735-934A-3
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: Li, Jiaxin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

```

```

Query Match      7.8%; Score 307.6; DB 4; Length 43950;
Best Local Similarity 67.6%; Pred. No. 3.4e-66;
Matches 491; Conservative 0; Mismatches 229; Indels 6; Gaps 4;

Qy 1442 TTTTTTTTTTTTTTTGAAGAAAGAGTCTTGCTCTTTACCCCTGGCTGGAGGCGAGTGTGT 1501
Db 9362 TTTTGTGTGTGTGAGACAGAGTCTTGCTCCCTCACCCAGGATGAGTGCATGGCGT 10021

Qy 1502 GATCTCAGCTCAGTGCACCTTTGCGCTCCCGGGTTCAAGCAATTTCTCTGCTTCAGGCTC 1561
Db 10022 GGTCTTGGCTCACTGCACAACTCTTACCTCTCTGGGTTCAAGCAATTTCTCTGCTTCAGGCTC 10081

Qy 1562 CCAAGTAGCTGGGACTACAGGCACTTCCCAACCATGGCCAGCTGATTTTGTATTTTAGT 1621
Db 10082 CCAATAGCTGGGATTAAGGTCGACCACTCTCTGGCTAATTTTGTATTTTAGT 10141

Qy 1622 AGAGATGGGATTTTCACTTTTGTGGCCAAAGTGTGCTCAAACTTTTGTCTGTCAATTTGT 1681
Db 10142 --AGACGGGTTTCAACATGTTGGCCAGGCTGTTTCAAGTAATCCACCTCTCAGCCT 10199

Qy 1682 TGTAACTATTGTTCTTTTGTCTGAGGTAGGGCCCCCAGACCAAAAAATAAATCTTTAGA 1741
Db 10200 CCCCAAGTGTGGGATTTACAGACATGAGCCACTGCGCTGGGCCCATATGCGCTCTCAATTT 10259

Qy 1742 ATCCAAATCAGTGTGTGTTTGTACCACTGCTCAGCTTGCAGAACCACTGTGACCGGGCC 1801
Db 10260 TGTATGTTGTGCCCTGCAATTAGAGCCATATTTCTTGATGTTTCAATTTGGGTATTAGGTCT 10319

Qy 1802 TCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAAAATAGAAATGAAATATTCTCGGGCCA 1861
Db 10320 GAGACAGCATCTTAGCTCCGTGGGTGCCACGCTTGTGACAGAAATCTCTGATTCTGGGCCA 10379

Qy 1862 GGGTGTGTGCTCATGCTCTTAATCCAGCACTTTCCAGGAGGCCCAAGGCAATGTGGATCACC 1921
Db 10380 GGACAGGTGTGCTCACCTGTAAATCCAGCACTTTGGGAGGCCCAAGGCGGCGGATCA-- 10437

```

QY 1922 TGAGGTCAAGAGTTCAAAACCCAGCTGGGCAACATGGTGAAACCCCGTCTCTACTAAAAA 1981
Db 10438 TGAGGTCAAGAGTTAGAGACCCAGCTGGGCAACATGGTGAAACCCCGTCTCTACTAAAAA 10497
QY 1982 TACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAATCCAGTACTTGGGAGGGTG 2041
Db 10498 TAGAAA-AAATTAGTGGTGTGGTGGCGGTACTATAATCCAGTACTTGGGAGGGCTG 10556
QY 2042 AGGCGAGGAGAAATTTCTTGAACCCG-GGAGCGAGAGGTTCAGTGAAGCGAGATCACCA 2100
Db 10557 AGGCGAGGAGATCAATTTGAACCTGAGGGGTGAGAGTTCAGTGAAGCGAGATCATACCA 10616
QY 2101 CTGCACTCCAGCTCGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 10617 TTGCACTCCAGCTGGGTGACAGGGTGAGACTCCGTCTCAAAAAACAAAAACAAAAAGA 10676
QY 2161 ATTAAG 2166
Db 10677 ATCCAG 10682

RESULT 12
US-10-060-332-3
; Sequence 3, Application US/10060332
; Patent No. 6528294
; GENERAL INFORMATION:
; APPLICANT: Li, Jiayin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL000851DIV
; CURRENT APPLICATION NUMBER: US/10/060,332
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-332-3

Query Match 7.8%; Score 307.6; DB 4; Length 43950;
Best Local Similarity 67.6%; Pred. No. 3.4e-66;
Matches 491; Conservative 0; Mismatches 229; Indels 6; Gaps 4;
QY 1442 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTACACCTGGCTGGAGGCGAGTGGT 1501
Db 9962 TTTTTTTGTGTGTGAGACAGATCTTGCTCGTCAACCCAGATGGAGTGGTGGT 10021
QY 1502 GATCTAGCTCACTGCAACTTTGGCTCCCGGTTCAAGCAATTCCTGCTTCAGCCTC 1561
Db 10022 GGTCTGGCTCACTGCAACTTACCTCTCGTGGTTCAGCAATTCCTGCTTCAGCCTC 10081
QY 1562 CCAAGTAGCTGGGATACAGGCACTTCCACCATGCCAGCTGATTTTGTATTTTGTAGT 1621
Db 10082 CCAATAGCTGGGATACAGTGGCGACCAACCACTCTCGCTGCTGCTGCTGCTGCTG 10141
QY 1622 AGAGATGGGATTCATTTTGTGGCAAGCTGGTCTCAAACTTTTGTGCTCATTAATGT 1681
Db 10142 --AGACGGGTTTACCATGTTGGCCAGGCTGTTTCAAGTAATCCACCTCTCAGCCT 10199
QY 1682 TGTAATATTGTTCTTTTGTGAGTAGGGGCCCCAGACCAAAAAAATAATCTTAGA 1741
Db 10200 CCCCAGGTGCTGGGATTAAGACATGAGCCATGCGGCTGGGCCCATCGCCTCATTTT 10259
QY 1742 ATCCAAATCAGTGTGTGGTTTGACCACTGTCTACTTGGAGCAACAGTGTGACAGGGCC 1801
Db 10260 TGATGTGTGCTCCGCAATTAGCCATATTCTTGGAGTTTCCATGGGTATTAGGTCT 10319
QY 1802 TCAGGAGTAGAGTGATCTCTCTCGAAAGAGAAATAGAAATGAAATATTTCTCGGGCCA 1861
Db 10320 GAGACGATCTCTAGTCCGTGGGTGCCACGCTGTGACAGAAATCCCTGATTTCTGGCCA 10379

QY 1862 GGCGTGGTGGCTCATGCCCTGTAAATCCAGCACTTTGGGAGGCCAAGGCAATGTGATCAC 1921
Db 10380 GGCAGGTGGCTCACACCTGTAAATCCAGCACTTTGGGAGGCCAAGGCGGGCGGATCA 10437
QY 1922 TGAGGTCAAGAGTTCAAAACCCAGCTGGGCAACATGGTGAAACCCCGTCTCTACTAAAA 1981
Db 10438 TGAGGTCAAGAGTTAGAGACCCAGCTGGGCAACATGGTGAAACCCCGTCTCTACTAAAA 10497
QY 1982 TACAAAAAATTAGCTAAGTGTGGTGGCGCATGCTCTAATCCAGTACTTGGGAGGGTG 2041
Db 10498 TAGAAA-AAATTAGTGGTGTGGTGGCGGTACTATAATCCAGTACTTGGGAGGGCTG 10556
QY 2042 AGGCGAGGAGAAATTTCTTGAACCCG-GGAGCGAGAGGTTCAGTGAAGCGAGATCACCA 2100
Db 10557 AGGCGAGGAGATCAATTTGAACCTGAGGGGTGAGAGTTCAGTGAAGCGAGATCATACCA 10616
QY 2101 CTGCACTCCAGCTCGGGGAGAGAGAGAGTTCCTCTCAAAAAACAAAAACAAAAAGA 2160
Db 10617 TTGCACTCCAGCTGGGTGACAGGGTGAGACTCCGTCTCAAAAAACAAAAACAAAAAGA 10676
QY 2161 ATTAAG 2166
Db 10677 ATCCAG 10682

RESULT 13
US-09-851-896-3/c
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851.896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match 7.8%; Score 307.2; DB 4; Length 70000;
Best Local Similarity 70.2%; Pred. No. 5.4e-66;
Matches 504; Conservative 0; Mismatches 193; Indels 21; Gaps 6;
QY 1441 TTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTTCAACCTGGCTGGAGGCGAGTGGT 1500
Db 50400 TTGTTTTTTTTTTTTTTGAGAGCGAGTCTCACTCTGTCAACCCAGGCTGGAGTGGCG 50341
QY 1501 TGATCTAGCTCACTGCAACTTTCGCTCCCGGTTCAAGCAATTCCTGCTTCAGCCT 1560
Db 50340 TGATCTAGCTCACTGCAACCTCTACCTACCGAGTTCAGCAATTCCTGCTTCAGCCT 50281
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCACCATGCCAGCTGA-TTTTGTATTTTGA 1619
Db 50280 CCAGATAGCTGGTATTACAGTCCCACTACGCCAGCTAAATTTTGTATTTTGA 50221
QY 1620 GTAGAGATGGGATTTCACTTTTGTGGCAAGCTGCTCAAACTTTTGTGCTGTATTAAT 1679
Db 50220 GTAGAGACAGGTTTCACTATTTGTCAGGCTGTCTTGAC---TCCTGACCTCGTG 50165
QY 1680 GTTGTAACTATTGTTTCTTTGCTGAGTAGGGCCCCAGACCAAAAAAATAAATCTTA 1739
Db 50164 ATTCCGCCACCTCGGCTCCCAAAATGCTGGATTACAGGCGTGAGCCACCGCGCCT-- 50107
QY 1740 GAATCCAAATCAGTGTGTGTTTGACCACTGTCTCACTTGAGAACCAAGTGTGACAGGG 1799

```

Db 50106 -----GCAITGGTTAAGCTTTTATGGTTGGCACT--GGATGGATGGATGGCCATTTA 50058
QY 1800 CCTCAGGAGTAGAGTGTGATCTCTGCTCGAAGAGAAATAGAAATGAAATATTTCTCCGGGC 1859
Db 50057 TGGTAGTGTGATGCAACATTTCTCTCAAAATTAATTAAGTCAATTTAAAGTAAGCCAGGC 49998
QY 1860 CAGCGTGGTGGTCTATGCTCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCA 1919
Db 49997 CGGCGAGGTGGTCTACGCTTGTAAATCCAGCAGTTTGGGAGGCTGAGGCAGGTGGAA 49938
QY 1920 CCTGAGTCAAGGAGTTCAAAACAGAGCTGGCCACATGTTGAAACCCGCTCTCTACTAAA 1979
Db 49937 CCTGAGTCAAGGAGTTCAAGACAGAGCTGGCCACATGTTGAAACCCGCTCTCTACTAAA 49878
QY 1980 AATACAAAAAATAGCTTAAGTGTGGTGGCGCATGCTGTAATCCAGCTACTTTGGGAGGG 2039
Db 49877 AATAC--AAATTAAGCAGAGTGTGGTGGCGCATGAACTTAATCCAGCTACTCGGAGGC 49820
QY 2040 TGAGGCAGGAGAAATTTCTGAACCCGGGAGCGAGAGTTGCGAGTGAAGCAGATCAACCC 2099
Db 49819 TGAGGCAGGAGAAATCGCTTGAACCCAGGAGCGAGAGTTGCGGTGAGCCAAAGATCGTGCC 49760
QY 2100 ACTGCACTCCAGCTGGGGA--GAGAGCGAGACTTCTCTCAAAAACAAAAACAA 2156
Db 49759 ATTGCACTCCAGCTGGGCAACAAAGCGAAACTCTCTCAAAATTAATAATAAAA 49702

RESULT 14
US-09-820-924-3/c
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
US-09-820-924-3

Query Match 7.7%; Score 304; DB 4; Length 39982;
Best Local Similarity 56.3%; Pred. No. 2.5e-65;
Matches 829; Conservative 0; Mismatches 595; Indels 48; Gaps 12;

QY 735 TATTTTCTTCTGAAACAGGCTATCACTCTGTTGCCAGGCTGGAGTACAGTGGCGTAA 794
Db 32734 TATTAATTTTGTAGATGAGTCTCACTTTGTGCGCCATGCTGGAGTGCAGTGCATGA 32675
QY 795 TAATAGTCACTCAGCCTCC--CCTCTGGGCTCAAGCAATCCGCTGGCTCAGCATCCT 853
Db 32674 TCTTGGCTCACTACAACCTCCACCTCCCGGTTCAAGCTATTCTTGTGCTCAGCCTCC 32615
QY 854 GAGTACTGGGACTACAGGCTTGTGCACAGGCCAGCTAAGTTTAAATAATGATTTT 913
Db 32614 GAGTACTGGGATTAAGGACCCACCATCACTGGCTTAATTTTT-----GTACTTT 32561
QY 914 TGGTATAGAGGAGTCTTGTATGTGCTCAGGCTGATTTTATTTTGTGAGACAAGGTC 973
Db 32560 TAGTAGAGATGGGTTCCACATGTTGGCGGCTGGCTTGAATCTGCTCAAGTGA 32501
QY 974 TCACTATGTTGCCATGATCCCCCACTCCACTTCCCAAGTGTCTCATC--TTATCTGTT 1031
Db 32500 TCCGCTGCTTTGGCTCCCAAGTGTGAGTTTACAGGTGTGTGCCACCAAGCCAGCT 32441
QY 1032 CATTAGTCACTGACAGACATTTAGGTTGTTTCCACTTTTGAACCAATTAATAATCT 1091

```

```

Db 32440 AATTTTGTATTTTATTAGTGAGAAGGGGTTTACCATGTTGGCCAGGCTGCTCGAACT 32381
QY 1092 CCAGTGAATATTCATGTATACATTTGTGTGGGCATATGTTTTCATTT----- 1138
Db 32380 CCTGACCTCTGTGATTCGGCGGCTCAGCCTCCCAAGTGTGGGATTAACAAGCCAGCC 32321
QY 1139 -----CTGTTGGGTTTATATCTAGGAGTGAATGTCTGGATCCCGGTAATATTTTGACA 1193
Db 32320 TGATACAGTGTGTTTAAATATGAGAGATCTCTACAAACCAATGTGGAATCATTTTCTTA 32261
QY 1194 GGCAGAGTTCCAGGGGAGAAAAAATTGGGAAAAATGAAGCATGTTTAG-----AAA 1243
Db 32260 ATAAATTTGTTAACTGAAGAAATGCAAAAGGAAAGACAGTAGTCTAGGGTACTACTTCTT 32201
QY 1244 TCAGCAGAGTGCAGGGGTTTTTCGGAGTTTTTATTTTATTTTCTGTTGACAAATGTGCAG 1303
Db 32200 GCTGTGAGAAAGGGAAGATAGCATACTTGTGTCTGCGCTGTTGTGCTAAACGAAC 32141
QY 1304 TTTGATGAAGATACAAAGTTATATCTAAGTGAAGATGAAGTTAAGGCTGGAATAGGGCGT 1363
Db 32140 CAGGAGGAGGAACCCAG--AAGTGAACGAGACTGCTTACCTACAAGGGGTAGGGGAGA 32083
QY 1364 TCAGAGTAAATCATGAGCAGCTTTGAATACCAAAATTAAGSAGCTTGGCTGTAAACAAA 1423
Db 32082 AAACAGGGCAAAAGGAGAGTTACTGTATTGTATAGTTCTGTCTTGTGGAGTACATTA 32023
QY 1424 ATAAATAAAATCACAAATTTTTTTTTTTTTTTTTTTGAGAAAGAGCTTGTCTTTTACCC 1483
Db 32022 TGTCTTTTTTTTTTTTTTTTATTTATTTATTTTATTTTATGACAGAGTCTCACTCTGT- 31964
QY 1484 GCTGAGGCGAGTGTGTGATCTCAGCTCACTCAACTTTTCGCTCC--GGTGTCAAGCA 1542
Db 31963 GCTGGAGTGCAGTGTGTGATCACAGCTCACTGAGCTCGACTCCAGGGCTCAGGTG 31904
QY 1543 ATTCTCTGCTCAGCTCCCAAGTGTGGGAGTGCAGGACTTCCCACCATGCCGAGC 1602
Db 31903 ATCTCCCACTCAGCTCCAGGAGTGTGGGAGTGTGAGTGTGCTACCAACCCAGC 31844
QY 1603 TGATTTTCTTATTTTATGAGAGTGGGATTTTCACTTTGTTGCCAAGCTGTGCTCAAC 1662
Db 31843 TAATTTTGTATTTCTCTGTAGACA--GAGGTTTGGCCATGTTGCCAGGCTGTCTCAAC 31785
QY 1663 T-TTTTGTCTCATAAATTTGTGTAACTATTGTTCTCTTTTGTGAGTGGGCCCCAGAC 1721
Db 31784 TCCTGGGCTCAAGCAATCCATCTGCTTGGCTCCCAAGTGTGATTTACAGGCATGA 31725
QY 1722 CAAAAAATAATCTTAGAATCCAAATCAGTGTGTGTTGACCTGTCTCATCTGAGA 1781
Db 31724 GACACTGTGCCAGCCACCGTTGTTTCAAAACATTCAAAAACAACAGCAACATAACAAAC 31665
QY 1782 ACCACAGTGTGACCCAGGCTCAGGAGTGTAGAGTGTATCTCTGCTCGAAAGAGAAATAGA 1841
Db 31664 CTCAAAACCAAAACAGAAATCAGCAAGATGAGAGGGAATCCCTAAGTGGAAATAAAA 31605
QY 1842 TGAATAATTTCTCCGGGCCAGGCGTGGTGGCTCAGCTGTAAATCCAGCACTTTGGGAG 1901
Db 31604 TGAATCCAA---CTAGGCTGGGCGGCTGGCTCAGACCTGTAAATCCAGCACTTTGGGAG 31548
QY 1902 GCCAAGGCATGTGGATCAGCTGAGGTCAGAGTTCAAAACCCAGCTGGCCAAACATGGTGA 1961
Db 31547 GCGAGGCGGGCGGATCAC--GAGGTCAGAGATCGAGACCATCTCTGGGTAAACCATGA 31490
QY 1962 AACCCCGCTCTCTACTAAAAATACAAAAATTAGCTAAGTGTGGTGGCGCATGCTGTAAT 2021
Db 31489 AACCCCATCTCTACTAAAAATACAAAAATTAGCCGGGCTGGTAGCGGCACCTGTAGT 31430
QY 2022 CCGAGCTATTGGGAGGTGAGCGAGGAAATTTCTTGAACCCGGAGGAGCAGAGTTGCA 2081
Db 31429 CCGAGCTACTCGGAGGCTGAGCGAGGAAATGGCGTGAACCCGGAGGCGAGAGCTTGCA 31370
QY 2082 GTGAAGCGAGATCACACACTGCTCCAGCTGGGGGAGAGCGAGACTTCTCTCTAA 2141
Db 31369 GTGAGCTGAGTAGTGCCACTGCATCTCCGCGCTGGGCGAAAGAGAGAGACTCCGTCTCAA 31310

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 09:06:53 ; Search time 2270.77 Seconds
(without alignments)

7909.918 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score: 3958

Sequence: 1 aggaattcatccattaaat.....gggacctcagcagccggg 3958

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3958	100.0	4150	14	US-10-016-725-15
2	3929.8	99.3	37698	13	Sequence 15, Appl
3	630.6	15.9	631	13	Sequence 10, Appl
4	630.6	15.9	631	16	US-10-021-632-24494
5	498.4	12.6	617	10	Sequence 24494, A
6	460	11.6	59725	13	US-09-792-4698-1
7	407.4	10.3	9631	15	US-10-087-192-814
8	360.4	9.1	23618	13	Sequence 814, App
9	353.6	8.9	36296	13	Sequence 1569, App
10	353	8.9	60815	13	Sequence 4, Appl
11	349.8	8.8	12394	16	US-10-240-425-1584
12	349	8.8	13216	10	Sequence 52, Appl
13	347.4	8.8	29163	10	US-10-087-192-52
14	342.8	8.7	26928	9	US-10-181-875-10

15	342.8	8.7	26928	13	US-10-235-192A-34	Sequence 34, Appl
16	342.8	8.7	26928	15	US-10-020-141-7	Sequence 7, Appl
17	342.8	8.7	26928	15	US-10-017-631-1	Sequence 1, Appl
18	338.6	8.6	144035	13	US-10-087-192-322	Sequence 322, App
19	337.2	8.5	3273	16	US-10-012-697-1406	GENERAL INFORMATI
20	337	8.5	109906	13	US-10-235-192A-31	Sequence 31, Appl
21	334.8	8.5	108182	13	US-10-087-192-1618	Sequence 1618, Ap
22	333.8	8.4	175077	13	US-10-087-192-1168	Sequence 1168, Ap
23	333.6	8.4	35425	15	US-10-017-161-2429	Sequence 2429, Ap
24	333.6	8.4	35425	16	US-10-292-798-2069	Sequence 2069, Ap
25	332.2	8.4	3158	13	US-10-027-632-115030	Sequence 115030, Ap
26	332.2	8.4	3158	16	US-10-027-632-115030	Sequence 115030, Ap
27	330.6	8.4	3158	13	US-10-027-632-115030	Sequence 115030, Ap
28	330.6	8.4	3158	16	US-10-027-632-115030	Sequence 115030, Ap
29	330.2	8.3	198522	13	US-10-087-192-244	Sequence 244, App
30	329.6	8.3	9372	9	US-09-764-877-3049	Sequence 3049, App
31	329.6	8.3	9372	9	US-09-764-877-3244	Sequence 3244, App
32	329.6	8.3	9372	16	US-10-242-515-3049	Sequence 3049, App
33	329.6	8.3	9372	16	US-10-242-515-3244	Sequence 3244, App
34	328.8	8.3	2130	16	US-10-108-260A-2246	Sequence 2246, App
35	328.8	8.3	75252	13	US-10-087-192-904	Sequence 904, App
36	328.6	8.3	35425	15	US-10-017-161-2429	Sequence 2429, App
37	328.6	8.3	35425	16	US-10-292-798-2069	Sequence 2069, App
38	328.4	8.3	96594	12	US-09-997-722-154	Sequence 154, App
39	328	8.3	91760	13	US-10-087-192-844	Sequence 844, App
40	327.6	8.3	12919	15	US-10-017-161-787	Sequence 787, App
41	327.6	8.3	17397	9	US-09-764-869-1945	Sequence 1945, App
42	327.6	8.3	17397	15	US-10-091-504-1945	Sequence 1945, App
43	327.6	8.3	17397	16	US-10-227-577-1945	Sequence 1945, App
44	327.6	8.3	19334	9	US-09-764-869-1943	Sequence 1943, App
45	327.6	8.3	19334	15	US-10-091-504-1943	Sequence 1943, App

ALIGNMENTS

RESULT 1

US-10-016-725-15
; Sequence 15, Application US/10016725
; Publication No. US20020151018A1
; GENERAL INFORMATION:
; APPLICANT: Prouty, Stephen
; APPLICANT: Zhang, Lin
; APPLICANT: Stenn, Kurt
; TITLE OF INVENTION: Stearyl-CoA Desaturase Promoter
; FILE REFERENCE: J&J2065
; CURRENT APPLICATION NUMBER: US/10/016,725
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 4150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-725-15

Query Match 100.0%; Score 3958; DB 14; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGGAATTCATCCATTTAAATCATACATTAATGCGTTTATAGTATATTCACAGGTTGTGC	60
Db	1	AGGAATTCATCCATTTAAATCATACATTTAAATGCGTTTATAGTATATTCACAGGTTGTGC	60
QY	61	ATCCATCAATCCATTTTATAGAACAGTTTATCTCCAAAAATAAACCTGCATTCCTT	120
Db	61	ATCCATCAATCCATTTTATAGAACAGTTTATCTCCAAAAATAAACCTGCATTCCTT	120
QY	121	AGCATACACCCCGAACATCTCCATCTCTCCAGCCCTGGGCAACCAACCAATCTAC	180
Db	121	AGCATACACCCCGAACATCTCCATCTCTCCAGCCCTGGGCAACCAACCAATCTAC	180
QY	181	TTTCTGCTCTATATAATTTGCAATTTCTGGACATTTTCATATAAATGGAAGCAACACAT	240


```
Db 631 AAGGTCCTCACTATGTTGCCATGATCCCGCCACCTCCACCTCCCAAGTGCATCTTATC 572
Qy 1028 TGTTCATTAGTCTAGTTCAGACAGACATTTAGGTTGTTTCCACTTTTTCACCAATTATGAATAA 1087
Db 571 TGTTCATTAGTCTAGTTCAGACAGACATTTAGGTTGTTTCCACTTTTTCACCAATTATGAATAA 512
Qy 1088 TACTCCAGTGAATATTCATGCTATACATTTGTGTGGGCATATGTTTTCATTTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATTCATGCTATACATTTGTGTGGGCATATGTTTTCATTTCTGTGGGT 452
Qy 1148 TTATATCTAGCAGTGGAAATGCTGGATCCCGGTAAATATTTTCACAGGAGAGTTCAGGG 1207
Db 451 TTATATCTAGCAGTGGAAATGCTGGATCCCGGTAAATATTTTCACAGGAGAGTTCAGGG 392
Qy 1208 GAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAAATCAGCAAGATGCAAGGGGTTTTTC 1267
Db 391 GAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAAATCAGCAAGATGCAAGGGGTTTTTC 332
Qy 1268 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTTCAGTGAAGATCAAGTTTATCT 1327
Db 331 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTTCAGTGAAGATCAAGTTTATCT 272
Qy 1328 AAGTGAAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 1387
Db 271 AAGTGAAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 212
Qy 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATATAAATAATCAAAATTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATATAAATAATCAAAATTTT 152
Qy 1448 TTTTATTTTGAAGAGAGTCTTGGCTTTTACCTGCTTTCAGGAGGAGTGTGATCTC 1507
Db 151 TTTTATTTTGAAGAGAGTCTTGGCTTTTACCTGCTTTCAGGAGGAGTGTGATCTC 92
Qy 1508 AGCTCACTGCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTTCAGGCTCCCAAGT 1567
Db 91 AGCTCACTGCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTTCAGGCTCCCAAGT 32
Qy 1568 AGCTGGGACTACAGGACATTTCCACCATGCC 1598
Db 31 AGCTGGGACTACAGGACATTTCCACCATGCC 1
```

RESULT 4

```
US-10-027-632-24494/c
; Sequence 24494, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24494
; LENGTH: 631
; TYPE: DNA
```

```
; ORGANISM: Human
US-10-027-632-24494
```

```
Query Match 15.9%; Score 630.6; DB 16; Length 631;
Best Local Similarity 99.8%; Pred. No. 1.7e-158;
Matches 630; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 968 AAGGTCCTCACTATGTTGCCATGATCCCGCCACCTCCACCTCCCAAGTGCATCTTATC 1027
Db 631 AAGGTCCTCACTATGTTGCCATGATCCCGCCACCTCCACCTCCCAAGTGCATCTTATC 572
Qy 1028 TGTTCATTAGTCTAGTTCAGACAGACATTTAGGTTGTTTCCACTTTTTCACCAATTATGAATAA 1087
Db 571 TGTTCATTAGTCTAGTTCAGACAGACATTTAGGTTGTTTCCACTTTTTCACCAATTATGAATAA 512
Qy 1088 TACTCCAGTGAATATTCATGCTATACATTTGTGTGGGCATATGTTTTCATTTCTGTGGGT 1147
Db 511 TACTCCAGTGAATATTCATGCTATACATTTGTGTGGGCATATGTTTTCATTTCTGTGGGT 452
Qy 1148 TTATATCTAGGAGTGGAAATGCTGGATCCCGGTAAATATTTTCACAGGAGAGTTCAGGG 1207
Db 451 TTATATCTAGGAGTGGAAATGCTGGATCCCGGTAAATATTTTCACAGGAGAGTTCAGGG 392
Qy 1208 GAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAAATCAGCAAGATGCAAGGGGTTTTTC 1267
Db 391 GAAGAAAACTTGGGAAAAATGAAGCATGTTTAGAAAATCAGCAAGATGCAAGGGGTTTTTC 332
Qy 1268 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTTCAGTGAAGATCAAGTTTATCT 1327
Db 331 GGAGTTTATTTTATATCTGTTGACAAATGTCAGTTTTCAGTGAAGATCAAGTTTATCT 272
Qy 1328 AAGTGAAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 1387
Db 271 AAGTGAAGTGAATTAAGGCTGGAATAGGCGTTTCAGAGTAAATCATGAAGCACCT 212
Qy 1388 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATATAAATAATCAAAATTTT 1447
Db 211 TGAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATATAAATAATCAAAATTTT 152
Qy 1448 TTTTATTTTGAAGAGAGTCTTGGCTTTTACCTGCTTTCAGGAGGAGTGTGATCTC 1507
Db 151 TTTTATTTTGAAGAGAGTCTTGGCTTTTACCTGCTTTCAGGAGGAGTGTGATCTC 92
Qy 1508 AGCTCACTGCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTTCAGGCTCCCAAGT 1567
Db 91 AGCTCACTGCAACTTTTCGCTCCCGGTTCAAGCAATTTCTCTGCTTTCAGGCTCCCAAGT 32
Qy 1568 AGCTGGGACTACAGGACATTTCCACCATGCC 1598
Db 31 AGCTGGGACTACAGGACATTTCCACCATGCC 1
```

RESULT 5

```
US-09-792-468B-1
; Sequence 1, Application US/09792468B
; Publication No. US20030157552A1
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brownlie, Alison J.
; APPLICANT: Ntambi, James M.
; APPLICANT: Miyazaki, Makoto
; APPLICANT: Gray-Keller, Mark P.
; APPLICANT: Attie, Alan D.
; TITLE OF INVENTION: Methods and Compositions Using Stearoyl-CoA Desaturase
; TO IDENTIFY TRIGLYCERIDE REDUCING THERAPEUTIC AGENTS
; FILE REFERENCE: 760050-7
; CURRENT FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-07-31
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: U.S. 60/184,526
; PRIOR APPLICATION NUMBER: U.S. 60/221,697
; PRIOR APPLICATION NUMBER: U.S. 60/255,771
```


Db 693 CTCAAGTATCTCTCTGCTCAGCTGCTTAATAGCTAGAGGACAGGGTGTGCCCAA 752
Qy 885 GGCCAGCTAAGTTTAAAAAATGATTTTGGTATAGAGAGGTCTTGTCTATGTGTCTCA 944
Db 753 TGCTAGCTAATTTTAAATTTTAAATTTTGGTATAGAGATGGATCTCACTATGTGCTA 808
Qy 945 GGCTGTATTTTATGTGTAGACAGAGTCTCACTATGTGTCATGATCCGCCACCTCCA 1004
Db 809 GGCTGCTGCTTGAATCTCTGGT-----TCAAGCAGTCTCTTGGCTCTCCGA 861
Qy 1005 CTTCCCAAGTCTCACTTATCTGTTCAATPAGTCACTGATGACAGACATTTAGTGTGTTTC 1064
Db 862 GTGTGTAGATGCGGGCGCAAGCACTCTGTCACCTGCTCTTCTTTTGTGTGATGTAG 921
Qy 1065 CACTTTTGGACCATTAATGAATPACTCCAGTGAATATCATGTATACATTTGTGTGGC 1124
Db 922 TAACTATTGATGATTTTACTTAGATCCATTAATCAATTAAGAGCTATACAATGATGATA 981
Qy 1125 ATATGTTTTCATTTCTGTTGGTTTATATCTAGAGTGGAAATGCTGGATCCCGGTAAAT 1184
Db 982 TTCTAGTCTTATTTATTTCTTTCAATTAATAGTCTGAATACTCTA-----TAAA 1029
Qy 1185 ATTTTGACAGCAGAGTTCAAGGGAAGAAAACCTTGGGAAATCAAGCANGTTTAGAAT 1244
Db 1030 GAGAACTCTCTTTCTAATCTTCGATTTACCCAGTGTATAATACATATTGGAAGT 1089
Qy 1245 CAGCAAGAGTGCAGGGGTTTTTGGAGTTTTTATTTATATTTCTGTTGACAAATGTGCAGT 1304
Db 1090 CAGGATGAATTTTGTCTTTCCCTGTTTACCAATTTTCAATATCTGATTTGGTTC 1149
Qy 1305 TTGATGAAGATACAAGTTATPACATGATGAGAGTGAATTAAGGCTGGAATAGGGGTT 1364
Db 1150 ACTAGCAATCTCCAA-TTAGGCATTTGTTTTGTGGGGTTTTTTTTGGTATCATAACTC 1208
Qy 1365 CAGATAAATCATGAAGCACTTTGAATACCAAAATTAAGAGCTTGGCTGTAACAAA 1424
Db 1209 TTGGGTTTAAATTTTGAATGATTTTCAATTCATTTGCAATTTTTTCTCTTTCTAGTGA 1268
Qy 1425 TAATAAAAA-----ATCAAAATTTTTTTTTTTTTTGGAGAAAGTCTGCTCTTTC 1477
Db 1269 TTTTACTGATATATTTCTTCTTCTTTCTTTTCTTTTGGAGCGAGTCTGCTCTGTT 1328
Qy 1478 ACCCTGCTGGAGGAGTGTGATCTCAGCTCACTGCAATTTGCGCTCCCGGGTTC 1537
Db 1329 GCCAGGCTGGAGTGCAGTGTGCAATCTCAGCTAACTGCAACTCGGCTCTTAGGGTTC 1388
Qy 1538 AAGCAATCTCTGCTTTCAGCTCCCAAGTCTGAGCTACAGGCACTTCCACCATGC 1597
Db 1389 AAGTGATTTCTCTGCTCAGTCTCTGATGATGGATTTACGCCATGCAACCATCGC 1448
Qy 1598 CCAGCTGATTTTGTATTTTATGATAGATGGGATTTCACTTTTGTGGCAAGCTGCTCT 1657
Db 1449 CCAACT-AAATTTGATTTTATGATAGATGGGTTTTCTCCATGTTTGTGTCAGCTGCT 1507
Qy 1658 CAAACTTTT-----TGCTGTCAATATTTGTGTAATCTATTTGCTTTGTGAGTGGGC 1713
Db 1508 CGAACTCTTTGACCTCAGGTGATCTCCCGCTCGGCTCCCAAGTACTAGGATTAACAG 1567
Qy 1714 CCCAGACAAAATAAATCTTAGAATCCAAATCAGTGTCTGCTGTGACCACTGTC 1773
Db 1568 CACCATGCCAGTGGCATATATATTTTCTTAATAAATCACTCAGCTTGATGATGAGCAGTG 1627
Qy 1774 ACTTGAGAACCAAGTGTGACCAAGGCTCTCAGAGTAGAGGTGATCTGCTGCAAGAGAG 1833
Db 1628 TCTTAGACATCTTCTGTCCATGCTGCTAGTGGCTGGCTTTTCAATAAATGCAATGAA 1687
Qy 1834 AATAGAAAT-----GAAATATTTCTCGGGCCAGGCGTGTGCTCA 1875
Db 1688 TGAGTTGATATTTCTGACGACGATTTAAAAATAAGAGCTAGTGTGCTGGCACGCTGCTCA 1747
Qy 1876 TGCTCTAATCCAGCACTTTGGGAGGCAAGGCAATGATGATCAGCTGAGGTGACAGGTT 1935
Db 1748 TGCTGTAAATCCAGCACTTTGGAGGCAAGGCGCGGATCAC--GAGGTGAGGAT 1805

RESULT 8

US-10-087-192-4/c
; Sequence 4, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 23618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23618)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-4

Query Match 9.1%; Score 360.4; DB 13; Length 23618;
Best Local Similarity 69.3%; Pred. No. 8.3e-85;

Matches 503; Conservative 0; Mismatches 221; Indels 2; Gaps 2;

Qy 1441 TTTTCTTTTCTTTTGTGAGAAAGAGTCTTGTCTTTTCCCTGGCTGGAGGCGAGTGGTG 1500
Db 11003 TTTCTTTTCTTTTGTGAGAGAGTCTGCTCTGTACCCAGGCTGGATGAGTGGCG 10944
Qy 1501 TGATCTCAGCTCAGTGCAACTTTGCGCTCCCGGTTTCAAGCAATTTCTCTGCTCAGCCT 1560
Db 10943 CGATCTCGGCTCACTGCAACCTTCACTCATAGGTACAGCAATTTCTCTCCCTCAGCCT 10884
Qy 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCACATATGCCAGCTGATTTTGTATTTTGG 1620
Db 10883 CCCGAGTAGCTGGGACTACAGGCGGCCCAACCATGCTGCTGCTGCTGCTGCTGCTG 10824
Qy 1621 TAGAGATGGGATTTTCACTTTTGTGGCAAGTGTGCTTCAAACTTTTGTGCTCATATTTG 1680
Db 10823 TAGAGACGGGTTTCAACATATTTGGCCAGGCTGGTCTGAACTCTCTGACTTGTGATCCG 10764
Qy 1681 TTGTAACTATTGTTCTCTTTTGTCTGAGTAGGAGGCCCCAGACCAAAAAATAAATCTTAG 1740
Db 10763 CCCACCTTGGGCTCCCAAGTGTGGGATTAAGGCTGAGGCTGAGGCTGAGGCTGAGT 10704
Qy 1741 AATCCAAATCAGTGTGTTGGTTT-TGACCACTGTCACTTGAGAACCAAGTGTGACGAGG 1799
Db 10703 TTAGCTATTCTTATTATTCAATTTATTTTTCAATTATGATCTACACACTGAGTTTATTA 10644

QY 1800 CCTCAGAGTAGAGGTGATCTCTGCTCGAAGAGAGAAATAGAAATATTTCTCGGGC 1859
DB 10643 CCTATGTATAAGATGTGTTATTTTTCGNNNNNNNNNNNNNNNNNNNNNTATATTTGGC 10584
QY 1860 CAGGGTGTGGCTCATGCTGTATATCCAGGACATTTGGGAGGCCAAGGATGTGGATCA 1919
DB 10593 CAGGCATGGTGGCTCATGCTGTATATCCAGCACTTTGGGAGGCTTGAGGCGAGCGATCA 10524
QY 1920 CCTGAGGTGAGAGTTCAAAACCCAGCCTGGCCCAACATGTAAGACCCCGTCTCTACTAAA 1979
DB 10523 CCTGAGGTGAGAGTTGAGATCAGCTGGCCCAACATGTAAGACCCCGTCTCTACTAAA 10464
QY 1980 AATACAAAAAATAGCTAAGTGTGGTGGCGCATGCTCTATATCCAGCTACTTTGGGAGGG 2039
DB 10463 AATAC-AAAAATTAACCGGGGTGGTGGCAGCGCTCTGTATATCCAGGTACTTGGGAGGC 10405
QY 2040 TGAGCAGAGAAATTTCTTGAAACCGGAGGAGGAGAGGTTCAGTGAAGCGGAGATCACACC 2099
DB 10404 TGAGCAGAGAAATCGCTTGTACCTTGAAGGAGGAGAGGTTCAGTGAAGCGGAGATCACACC 10345
QY 2100 ACTGCACCTCCAGCCTGGGGGAGAGAGAGCTTCTCTCAAAAAAACAACAAAG 2159
DB 10344 ATTGCACCTCCAGCCTGGGTGACAGAGCAGACTCTGTGCGCAAAAAAATAATTATA 10285
QY 2160 AATTAA 2165
DB 10284 ATTTTA 10279

RESULT 9
; Sequence 1584, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1584
; LENGTH: 36296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 Z82180
US-10-240-425-1584

Query Match 8.9%; Score 353.6; DB 13; Length 36296;
Best Local Similarity 57.2%; Pred. No. 7.5e-83;
Matches 956; Conservative 0; Mismatches 609; Indels 107; Gaps 13;

QY 522 GCTAATTTTTTTTTTAATTTTCAATTTATGTAATTCATTTCTTTCTTTTGTGTGT 581
DB 15559 GGTAATGTTAGCTTTCTCAAGCAGCATTAATATCAATTTATTTATCAATGTAATTT 15500
QY 582 TGTTTTGAATAGGTCTCACTTTGTTTACCAGGCTGGAGGCGAGTGGCATGGTGACAGC 641
DB 15499 TTTTATAGAGCAAGGTCTCAATCTCCACCCAGGCTGGAGTGTGTGCAATCATAGCTC 15440
QY 642 TGAGCAGCCTTGACTTCTCTGGGCTCAAGTGATCTCTCTGCTCAGCTCCCAAGTAGCTG 701

DB 15439 ACTGAGCCTCCAACTCTCTGGGCTCAAGTGATCTCAGCTCCTGAGTAGTTG 15380
QY 702 GGACTACAAACACATGTCACCATGCTGGCTGATATTTTTTTTCTTGAACAGGGTATCA 761
DB 15379 GGACTGCAAGTGCATGCCCATGCTGGCTTTTTTTTTTTTTCAGACAGAGTTTCGCT 15320
QY 762 CTCTGTTGCCAGGCTGGAGTACAGTGGCGTAATAATAGCTCACTGCAAGCTCC-CCTCC 820
DB 15319 CT-TGTCACCCAAGCTGGAGTGCAGTGGCAGATCTTGGCTCACTGCAACCTCGCTCC 15261
QY 821 TGGCTCAAGCAATCCGCTGGCTCAGCATCTCAGTAGTGGAGCTACAGGCTGTGTC 880
DB 15260 CAGGTTCAAGTGATCTCTCGCTCAGCTCCTGAATAGCTGGGATTACAGGCAACACC 15201
QY 881 ACCAGGCCACGCTAAGTTTTTAAAAATGATTTTTTGTATAGAGGAGGCTTTCCTATGTTG 940
DB 15200 ACCAGGCTGTCTAATTTTT-TGTATTTTAGTAGACAGGGTTTCAACATGTTG 15146
QY 941 CTCAGGCTGTATTTTTTATTTGTAGACAAAGTCTCACTATTTGCCATGATCCGCCACC 1000
DB 15145 GGCAAGCTGGTCTCTAACTCTGACCTCAGG-----TGATCCGCTGCC 15102
QY 1001 TCCATTTCCCAAGTGTCTCATCTTATCTGTTCATTAGTTCAGTTCAGACAGATTTAGTTG 1060
DB 15101 TCAGCCTCCCAAGTGTCTGGGATTACAGGTGTAGGGGCCATTT-TTTTTTAAAGTAC 15044
QY 1061 TTTCCACTTTTGTACCAATTAATAATATCTCCAGTGAATATTCATGTATATACATTTGTGT 1120
DB 15043 AGATGGGCTCTTGCTATGTTGCTCAGGCTGGTCTCAAAATCTTTGGGCTCAAGTATCCTC 14984
QY 1121 GGGCATATGTTTTTCATTTCTGTGGGTTTATATCTAGAGTGGAAATTGCTGGATCCCGGG 1180
DB 14983 CCTCTGAGCCTCCCAAGTGTGGGATTATAGGACAGGACACCGCACCCAGCTCAGG 14924
QY 1181 TAATATTTTACAGGCGAGGTTTCAAGGGGAGAAAACTTGGGAAAAATGAACATGTTTAG 1240
DB 14923 TGACATTTTAATGGGAGAGGANGTGTGTGAGTGCATCCAAAGGGACAGCTCCTGATG 14864
QY 1241 AAATCAGCAAGTGCAGGGGTTTTTCGGAGTTTATTTTATATTTCTGTGACAAAATGTG 1300
DB 14863 CTGTTAGAGCCAGCTAGAGTTTGGTCTTCCCTTAGTGTGAGTTAGGAGAGATTTTT 14804
QY 1301 CAGTTTGTATGAAGATCAAGTTTATCTAAGTGAGAGTGAGAAATTAAGGCTGGAAATAGG 1360
DB 14803 CTGTGCCAAGAGCCCTGCAGTT-----14782
QY 1361 CGTTCAGAGTAAATCATGAAGCACTTTGAATACCAAAATTAAGGAGCTTGGCTGTAAAC 1420
DB 14781 -----CAGTGGTGGGACCTGACAACTCTCTTATAGGAAGACACC 14741
QY 1421 AAAATAAATAAATAACAAATTTTTTTTTTTTTTTTGGAGAAAGAGTCTTGTCTTTCAAC 1480
DB 14740 CTCCTGTAGTCTTTTTTTTTTTTTTTTTTTTCTTTTTCAGACAAAGTCTTGTCTTTCA-C 14682
QY 1481 CTGCTGGAGGCGAGTGTGTGATCTCAGCTCACTGCAACTTTTGCCTCCCGGGTTCAG 1540
DB 14681 CAGGCTGGAGTGCAATGGCATGATCTCGGTCTCACTGCAACCTCCACCTCTTGGGTTCAAG 14622
QY 1541 CAATTCCTCTGCTCAGCTCCCAAGTAGCTGGAGCTACAGGCACTTCCCAACCATGCCA 1600
DB 14621 CGATTTCTTCTGCTCAGCTCCCGAGTAGTGGGAGTACAGGCAACCGCCACCAATCTG 14562
QY 1601 GCTGATTTTTGTATTTTTTAGTAGAGTGGGATTTCACTTTTGTGGCAAGCTGTCTCAA 1660
DB 14561 GCTAAATTTTGTATTTTTTANTAGAGGGGAGTTCACCATGTTGGCCAGGATGGTCTCGA 14502
QY 1661 ACTTTTGTGTGATTAATTTGTGTAATATGTTTCC-----TTTTGTGTAGTAGGGC 1713
DB 14501 TCTCCTGACCTATGATCTGCTGCTCGGCTTACAGGCATGAGCCACTGCGCCCGGCC 14442
QY 1714 CCCCAGACCAAAAAATAAATCTTAGAATCCAAATCAGTGTGTGTTGTGACCACTGTC 1773

Qy	964	---	AGA	CAGGTCTCACTATGTTGGCCATGATCCCCACCTCCACATTCCCAAAGTCTCTC	1019
Db	4195	TTTT	AA	GACAGAAATCTCGTTCTGTCTCACTAAGCGTGGAGTGCAGTGGTGTGATGTCGGCTCA	4254
Qy	1020	ATCT	TATCTGTTCATTAGTCTAGTTGACAGACATTTAGTGTGTTTCCACTTTTTCACCAATT	1079	
Db	4255	CTGC	AACTCTCGCTCCCTGGGGTTCAAGCGCTGCACCTCTCTCCCTCTCTGGGTTCAGGCAAT	4314	
Qy	1080	ATGA	ATAATCTCCAGTGAATATTCATGTATATACATTTGTGTGGGCATATGTTTTCATTTC	1139	
Db	4315	TCTT	GTACTCATCCACTGAGTAGTTGGAATCATACAGSCGTGCGCCACCATGCCACGTA	4374	
Qy	1140	TGTT	GGGTTTATATCTAGAGTGGAAATTCGTGATCCCGGGTAATATTTTGCACAGGCGA	1199	
Db	4375	ATTT	TTTTTTGTATTTTATAGTAGATGGGGTTTTCGCCAGTTGGCCAGGCTGGTCTCGAAC	4434	
Qy	1200	GTTC	AGGGGGAAGAAA-----	1246	
Db	4435	TCCT	GGCCCTCAAGTGAATTCCTCGCTTGGCCTCCCAAAGTCTGGGATTCAGGGGGTGA	4494	
Qy	1247	GCA	AGAGTCAGGGGTTTTTTCGGAGTTTTATTTTTATATCTCTGTGTGACAAATGTGCAGTTT	1306	
Db	4495	GCC	ACCATGCCAGCTGTTTTTATTTTTATTTTTATTTAAG--GCTGGGTATGTTAGCT	4552	
Qy	1307	GAT	GAAGATACAGATTATCTAAGTGAGAGTGAGAAATTAGGCTGGAAATAGGGCGTTCA	1366	
Db	4553	CAT	CGCTGTATCTCTTGAATCTTGGAGGCC--CGAGGCGAGGAGGATTCCTCTGAGACTAGGA	4611	
Qy	1367	GAG	TAAATCATGAAGCACATTGTGAATACCAAATTAAGGAGCTTGGCTGTAAACAAATA	1426	
Db	4612	GTT	CAAAACCACTGGCCACATAGCCAGGTTCTTTTAAATAATAATAATAATAAT	4671	
Qy	1427	AT	AAAAATCACAAATTTTTTTTTTTTTTTTGGAGAAAGACTTGTCTTTTCAACCTCGGCT	1486	
Db	4672	TTT	ATCTTATTTATTTATTTATTTATTTTGGACAGAGTCTGTCTGCGCCAGGCT	4731	
Qy	1487	GG	AGGAGTGGTGATCTCAGCTACTGTGCACTTTTCGCTCCCGGTTTCAGCAAAATTC	1546	
Db	4732	GG	AGTGCAGTGGCGCGATCTCAGCTCACTGCAAGCTCCGCTCCTGGTTTCACGCATTC	4791	
Qy	1547	TCC	TGCTTCAGCCFCCCAAGTAGCTGGAGTCTCAGGCACTTCCCAACATGCCAGCTGA-	1605	
Db	4792	TCT	TGCTCAGCTCCCGAGTAGCTGGAGTACAGTGCCTGCGCCATGCTTGGCTAAT	4851	
Qy	1606	--	TTTT	TGTAATTTTAGTAGAGATGGGATTCATTTTGTGGCCAAAGCTGGTCTCAAACT	1663
Db	4852	TTTT	TTTT	TGTAATTTTAGTAGAGACAGGGTTTACCGGTGTTTAAACAGGATGGTCTCACTCT	4911
Qy	1664	TTTT	TGCTGTATATTTGTTGTAATAATTTGTTCTTTTGTGAGTGGGCCCCGACCA	1723	
Db	4912	CT	GTAC-----	TTGTTGATCCACCACTCAGGCTCCCAAGTGTGGGATTCAGGCGT	4966
Qy	1724	AAAA	AAAAATTAATCTTAGAATCCAAATCAGTGTGTGTGTTTGACCACTGTCACTTGGAGAC	1783	
Db	4967	GAG	CCACCGCTTGGCCCT-----	GGCTATTCCTTTTAAACTTTATTTTGGAGAA	5019
Qy	1784	CAC	AGTGTACCAAGGCTCTCAGGAGTAGAGTGATCTGTCTCGAAAGAGAAATAGAAATG	1843	
Db	5020	AAAA	ATCAGAAGTGCATTTGGCTTTTACATGTCAGCAATAAGTTGAAAAAATTTT	5079	
Qy	1844	AAAA	ATATCTCCGGGCCAGCGGTGGTGTCTATGCTGTAAATCCAGCACTTTGGAGGC	1903	
Db	5080	TTTT	TAAGTGGGGTGGTGGGCGGGTGCCTTCAGCCTGTAAATCCAGCACTTTGGAGGC	5139	
Qy	1904	CAG	GCATGTGGATCA-----	CCTGAGTCTAGGAGT	1934
Db	5140	TG	AGCCTGTGGATCATGAGGTCAAGGAGGCTGAGCGAGTGGATCACAAAGTTCAGGAGA	5199	
Qy	1935	TCA	AAACCGCCTGGCCCAACATGTTGAACCCCGCTCTCTACTAAATAATCAAAAAATTAG	1994	
Db	5200	TC	GAGCACTCTTGGCTTAAGTGGTGAACCCCATCTCTACTTAAATATAC--AAAAATTAG	5258	
Qy	1995	CT	AGTGTGGTGGCGATGCTCTGTAAATCCAGCTACTTGGGAGGGTGCAGGCAAGGAAT	2054	

Db	5259	CTGGGCGCTGGTGGTCATCGCTGTAATCCAGCTACTTGGAAAGCTGAGGCAGGAGAA	5318
QY	2055	TCTTTGAACCGGGAGGAGGAGGTTGTCAGTGAAGCAGATCACACCAC	2114
Db	5319	GCTTTGACCCCGGAGGAGGAGGTTGTCAGTGAAGCAGGATATTGAGC	5378
QY	2115	GGGGGAGAGGAGGAGACTTCTCTCAAAAAAACAAAAA	2153
Db	5379	GGCAACAGAGCAGAGACTCTGTCTCAAAAAAACAAAAA	5416

RESULT 12

US-09-764-891-7810

; Sequence 7810, Application US/09764891

; Publication No. US20030077808A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC006

; CURRENT APPLICATION NUMBER: US/09/764,891

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 10231

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7810

; LENGTH: 13216

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-891-7810

Query Match	8.8%;	Score 349;	DB 10;	Length 13216;
Best Local Similarity	70.0%;	Pred. No. 6.5e-82;		
Matches 516;	Conservative 0;	Mismatches 210;	Indels 11;	Gaps 3;
Qy	1441	TTTTTTTTTTTTTTTTTGAGAAAGAGTCTTGCTCTTTCCACCTCGCTCGAGGGCAGTGGTG	1500	
Db	3625	TCYTTTCYTTTTTTTTTTCGAGTCTTGCTTTGTGCGCCAGGCTGGAGTGCAGTGGTG	3684	
Qy	1501	TGATCTCAGCTCACTGCAACTTTTCGCTCCCGGGTTCAAGCAATTCCTCGTTCACGCT	1560	
Db	3685	CGATCTCAGCTCACTGCAAACTCCACTCCGCGGGTTCAAGAGATTCCTCGCTTAGCCT	3744	
Qy	1561	CCCAAGTAGCTGGGACTACAGGCACATCCACCACATGCCAGCTGATTTTTGTATTTTTAG	1620	
Db	3745	CCTGAGTAGCTGGGATTACAGGGGTGCATACCTGTCGCTGGCTAATTTTTTATTTTTTAA	3804	
Qy	1621	TAGAGATGGGATTCTCATTTTGTGGCGAAGCTGGTCTCAAACTTTTTCGTGTCATAATTG	1680	
Db	3805	TACAGATGGGGTTTCACCATGTTGGTTCAGGCTGGTCTCGAACTTCCTGACTTCATGATCCG	3864	
Qy	1681	TTGTAACTATTGTTTCCTTTTGTCTGAGTAGGGCCCCAGACCAAAAAATAAATCTTAG	1740	
Db	3865	TCGCCCTCGACCTCCCAAGTGTAGGATTACAGGCATGAGCCACTGTGCCAGCC---	3920	
Qy	1741	AATCCAAATCAGTGTGTTGGTTTGACCACTGTCACTTGAGAAACCAAGTGTGACCAAGGC	1800	
Db	3921	---GCTAAATTTATTTTGTACAGACAGAGTTTCTCCATGTTGCCAGGCTGATCTGGAA	3977	
Qy	1801	CTCAGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATGAATAATTCCTCCGGCC	1860	
Db	3978	CTCTGGGGCGAAGTGAAACATCTGCTCGCC---CTCCAAAGTGTGGGATTACAGGGC	4034	
Qy	1861	AGCGGTGGTCTATGCTGTAATCCACGACTTTTGGAGGCCAAGGATGTGGATCAC	1920	
Db	4035	GGGTGGGTGGTCTACGCTGTAAATTCAGACACTTTGGGAGGCCAAGCGAGTGGATCAG	4094	
Qy	1921	CTGAGTCCAGGAGTTCAAACACAGCTGCGCAACTGGTGAACCCCGTCTCTACTAAAA	1980	
Db	4095	CTGAGTCCGGAGTTCAAGACCAGCTGGCCCAACAGGTGACACCTGTCTCTACTAAAA	4154	
Qy	1981	ATACAAAAATTAAGCTAAGTGTGGTGGCGCATCGCTGTAAATCCACGCTACTTGGGAGGCT	2040	

Db	4155	ATAC - AAAAATTAGCCACGTTGGTGGCGCATGCCTATAGTCCCACTACTTGAGAGGCT	4213
Qy	2041	GAGCGAGGAGAAATTTCTTTGAACCCGCGAGGCAGAGGTTGCAGTGAAGCGAGATCACAACCA	2100
Db	4214	GAGCGAGGAGAAATCATTGAACCCGCGAGGCAGAGGTTGCAGTGAAGCGAGATCACAACCA	4273
Qy	2101	CTGCACATCCAGCCTGGGGGAGAGAGCAGTCTCTCTCAAAAAACAACCAAAACAAGA	2160
Db	4274	CTGCACATCCAGCCTGGGGGAGAGAGTGTGTCTCAAAAAACAACCAAAACAACA	4333
Qy	2161	ATTAAGCAAAATTAGACA	2177
Db	4334	ACAAAACAACAAACA	4350

RESULT 13

US-09-764-891-7809

; Sequence 7809, Application US/09764891

; Publication No. US20030077808A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC006

; CURRENT APPLICATION NUMBER: US/09/764,891

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 10231

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7809

; LENGTH: 29163

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-891-7809

Query Match	8.8%;	Score 347.4;	DB 10;	Length 29163;
Best Local Similarity	69.9%;	Pred. No. 3e-81;		
Matches 515;	Conservative 0;	Mismatches 211;	Indels 11;	Gaps 3;
QY	1441	TTTTTTTTTTTTTTTTTGAGAAAGAGCTTGCTCTTTTCAACCTGGCTGGCTGAGGGGAGTGGTG	1500	
Db	3623	TCCTTTCTTTTTTTTTTTCTGGAGCTTGTGCTTGTGCGCCAGCGTGGAGTGCAGTGGTG	3682	
QY	1501	TGATCTCAGCTCACTGCAACTTTTGGCTCCCGGGTTCAAGCAANTTCTCTGCTTCAAGCCT	1560	
Db	3683	CGATCTCAGCTCACTGCAAACTCCACTCCCGGGTTCAAGAGATTCTCTCTGCTTAGCCT	3742	
QY	1561	CCCAAGTACGTGGAGCTACAGGCACATCCACCATGCCAGCTGATTTTGTATTTTAT	1620	
Db	3743	CTGTAGTACGTGGATTACAGGGGTGCATACCCTGCTGGCTAAATTTTATATTTTAA	3802	
QY	1621	TAGAGATGGGATTTCACTTTTGTGGCAAGCTGGTCTCAAACTTTTGTCTCATATG	1680	
Db	3803	TACAGATGGGTTTCAACATGTGGTCAGGCTGGTCTCGAACTTCCTGATCTCATGTCG	3862	
QY	1681	TTGTAACTATTTGTTCTTTTGTAGGTAGGGCCCCCAGACCAGCAAAAAAATAAATCTTAT	1740	
Db	3863	TCCCCCTCGACCTCCCAAAGTGTTAGGATTACAGGCATGAGCCACTGCCCCAGCC---	3918	
QY	1741	AATCCAAATCAGTGTGTTGGTTGACCATGTCTCATTTGAGACCAAGTGTGACACGGGC	1800	
Db	3919	---GCTAAATTAATTTTGTACAGACAGAGTTTCTCCATGTTGCCAGGCTGATCTGGAA	3975	
QY	1801	CTCAGGAGTAGGGTGATCTCTGCTGMAAGAGAAATAGATGAAAATATTTCTCGGGGC	1860	
Db	3976	CTCTGGGGCAGTGGACCAATCTGCTCGGC---CTCCAAAGTGTGGGATTAAGGCC	4032	
QY	1861	AGCGTGGTGGCTCATGCTTAATCCAGCATTTTGGGAGGCCAAGGCATGTGGATCAC	1920	
Db	4033	GGGTGCGGTGGCTCAGCGCTGTAATTCAGCATTTTGGGAGGCCAAGCGAGTGGATCAG	4092	
QY	1921	CTGAGTCAAGGATTCAAAACCAAGCTGGGCCAATGTTGGTGAACCCCGTCTCTACTAAA	1980	
Db	4093	CTGAGTCCGGAGTTCAAGACAGCTGGGCCAAGGTTGACACCTGTCTCTACTAAA	4152	

QY	1991	ATACAAAAAATTAGCTAAGTGTGTGTGGCGCATGCTCTTAATCCAGACTACTTTGGGAGGGT	2041
DB	4153	ATAC-AAAAAATAAGCCACGCTGTGTGTGGCGCATGCTCTAATAGTCTCCCACTACTTGTGAGAGGCT	4211
QY	2041	GAGCGAGGAGAAATTTCTTTGAACCGGAGGAGCAGAGGTTGCAGTGAAGCGAGATCACAACA	2100
DB	4212	GAGCGAGGAGNATCACTTTGAACCGGAGGAGGAGGTTGCAGTGAAGCGAGATCACAACA	4271
QY	2101	CTGCACCTCCAGCTGTGGGGAGAGGCGAGACTTCTCTCTCAAAAAACAAAAACAAAAGA	2160
DB	4272	CTGCACCTCCAGCTGTGGGGAGAGGAGTGTGTCTCTCAAAAAACAAAAACAAAAGA	4331
QY	2161	ATTAAGCAAAATTAGACA	2177
DB	4332	ACAAAACAACAACA	4348
RESULT 14			
US-09-880-107-2278			
; Sequence 2278, Application US/09880107			
; Patent No. US20020142981A1			
; GENERAL INFORMATION:			
; APPLICANT: Horne, Darci T.			
; APPLICANT: Vockley, Joseph G.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Gene Logic, Inc.			
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
; FILE REFERENCE: 44921-5028-WO			
; CURRENT APPLICATION NUMBER: US/09/880,107			
; CURRENT FILING DATE: 2001-06-14			
; PRIOR APPLICATION NUMBER: US 60/211,379			
; PRIOR FILING DATE: 2000-06-14			
; PRIOR APPLICATION NUMBER: US 60/237,054			
; PRIOR FILING DATE: 2000-10-02			
; NUMBER OF SEQ ID NOS: 3950			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2278			
; LENGTH: 26928			
; TYPE: DNA			
; ORGANISM: Homo sapiens.			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M17262			
US-09-880-107-2278			
Query Match 8.7%; Score 342.8; DB 9; Length 26928;			
Best Local Similarity 57.5%; Pred. No. 4.9e-80;			
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;			
QY	735	TATTTTCTTTTGTGAACAGGGATACATCTGTGTGGCCAGGCTGGAGTACAGTGGCGTAA	794
DB	3122	TTTCTTTTGTGGAGTCAAGAGCTTGTCTGTGTGGCCAGGCTGGAGTACAGTGAACGA	3181
QY	795	TAATAGTCACTCGACGCTCC-CCTCCTGGCTCAAGCAATCGCTGGCTCAGCATCT	853
DB	3182	TCTCAGCTCACTCGACGCTCCGCTCCGCGTCAAGCTATTCTCTGTCTCAGGCTCCC	3241
QY	854	GAGTAGCTGGGACTACAGGCTTGTGCGCACACAGCGCCAGCTAAGTTTTTAAAAAATGATTTT	913
DB	3242	AAATAGCTGAGACTATAGGACGCGACATCCATGCCAGCTAATTTT-----TTATTTT	3295
QY	914	TGCTATAGAGGAGGCTTCGTAATGTGTCTCAGGCTGTATTTTATTTGTGAGACAAGGTC	973
DB	3296	TAGTAGAGACGAGGCTGTCTCCATGTGGCGCAGTGTGGTCTTGAACCT-----3341	
QY	974	TCACTATGTTGCCATGATCCCCCACCCTCCATCTCCAAAGTGCTCATCTTATCTGTTCA	1033
DB	3342	--CCTGCTCAAGTGATCCACTGCGCTCGGCTCCAAAGTGGGATTCGAGGCATG	3399
QY	1034	TTAGTCAGTTGACAGACATTTAGTTGTTTCCATCTTTTGACCATTAATGAATAATACTCC	1093
DB	3400	AGACACCGCGCCCGGCC-----TGCTGTGCTCCCTTCTTAAATGAGTTGTCCATTGTAA	3454

Qy	1094	AGTGAATATTCAATGTAATACATTTGTGTGGCGAATATGTTTTCAATTTCTGTGTGGGTTTATAT	1155
Db	3455	GCTGCTGATTTCTTTTGGGACATTTGCTCTCGTAAACTTTTTCATTAAGCATCAGTGAATTC	3514
Qy	1154	CTAGGAGTGGAAATTTGCTCGATCCCGGGTAAATATTTTGACAGCGACAGCTTCAGGGGAAGAA	1213
Db	3515	CCATTCCTTCCACCCAGCTTCACCGTAATTTGTGTGTGTCTTCAATTTTCAGTTCAGCA	3574
Qy	1214	AAACTTTGGGAAATTAAGACATGTTTGAATAATCAGCAAGAGTGCAGGGGTTTTTTCGAGATT	1273
Db	3575	GAAATTCATTTAGCTCTGATAAGGCGTCGCTCAAACTGATGCTCTTATCTCTTCTAGTGCC	3634
Qy	1274	TTATTTTAAATCTTTGTGACAAATGTGC-----AGTTTGATGAAGATACAAAGTTTATACTA	1328
Db	3635	TCAAACACTACATCTCTGTTTCACTCATGTTATAGCAAGTTAGTGTGAGTTATTTTGGTGCAC	3694
Qy	1329	AGTGAGAAGTGAAGTAATTAAGCTCGAATAGGGCGTTCAAGAGTAAATCATGAGACACTTT	1388
Db	3695	AAAAATTTTTTAAATCCATGCACTCTTTTTTTCATAATACGCATTTTCCATGAACTTTTC	3754
Qy	1389	GAATACCAAAATTAAGGAGCTTGCGCTGTAAACAAAAATAAATAAAAAATCACAATTTTTTTTT	1448
Db	3755	GAAGACCCCTGTGATAGTCTGTGTGTTTAAACACCCAGTTTA-----CAGTAATTTTTTTT	3810
Qy	1449	TTTTTTTTTGAGAAAGAGTCTTTGCTCTTTTCACTTCCCTGGCTGGAGGCGAGTGTGTGATCTCA	1508
Db	3811	TTTTTTTTTGAGATGAAGTCTTTGCTCTCTCGCCAGGCTGGAGTGCATTTGGCAACACTCTCG	3870
Qy	1509	GCTCACATGCAACTTTTCGCCCTCCGGGTTCAAGCAATTTCTCCTGCTTCAGCGCTCCCAAGTA	1568
Db	3871	GCTCACATGCAACCTTCTGCCTCCTGGGTTCAAGCAATTTTCTGCTCAGTCTCCCGAGTA	3930
Qy	1569	GCTGGGACTACAGGCACTTCCACCATCGCCAGCTGATTTTTTGTATTTTTTAGTAGAGATG	1628
Db	3931	GCTGGGATTACAGTGTGTGCCACCATGCTCTAGCTAATTTATGTGTTTTTTAGTAGAGCG	3990
Qy	1629	GGATTTCACTTTTGTGTGCCAAGTGTGTTCAAACTTTTTGCTGTCAATATTTGTTGTAACT	1688
Db	3991	GGGTTTTCACTATGTTGGCTAGGCTGTGCTCGAACTCCTCACTTGATCGGCCCGCCTC	4050
Qy	1689	ATTGTTTCCTTTTGTGTAGGTAGGCCCCGAGCAAAAAAATAAATCTTTAGAAATCCAAA	1748
Db	4051	GGCCTCCCAAAGTATCGGGAATTACAGCGGTGA-----GACTCTTGCATCTGGCC	4099
Qy	1749	TCAGTGTGTGTGGTTTGACCACCTGTCACTTCAGAAACACACAGTGTGACACAGGCGCTCAGGAG	1808
Db	4100	TACAGTAATTTTATAGCAGCTTAGCTTAAGTAGCCATTTCTGGGTATAGAAATGTCTATA	4159
Qy	1809	TAGAGGTGATCTCTGTCTGAAAGAGAATAAGAAATGAAATAATTTCTCGGCGCCAGGCGTGG	1868
Db	4160	TACTGAACAGCGCTCGCAACTGTGAGTAAAAAGTCTGCAAGAGGCC-----GGGCAG	4210
Qy	1869	TGGCTCATGCTGTAAATCCAGACACTTTGGAGGCCAAGGCATGTGAGTCACTGAGGCTC	1928
Db	4211	TGGCTCATACTGTAAATCCAGACACTTTGGGGGCGAGCAGGTGGATCACCTGAGGCTC	4270
Qy	1929	AGGAGTTTCAAAAACAGCGCTGGCCAAATGTTGAAACCCCGCTCTCTACTAAAAATACAAAA	1988
Db	4271	AGCAGTTTCGAGACAGCCTGACACATGTTGAAACCCCATCTCTACTAAAAATATC--AA	4328
Qy	1989	AATPAGCTAATGTGTGGCGCATGCTGTAATCCC-----AGCTACTTGT	2033
Db	4329	AATTAGCTGGGCGTGGTAGTGATGCTTTGATTCCTCTAGCATGCATTTGGGAGCTACTTGT	4388
Qy	2034	GGAGGTTGAGCAGGAGAAATTTCTTGAAACCGGGAGGCAGAGTTGCATGTAAGCCAGAT	2093
Db	4389	GGAGGCTTGAGCAGGAGAAATCACTTTGTACTCAGGAGGCGAGGTTGCAGTGCAGTGAGAT	4448
Qy	2094	CACACCACTGCATCTCAGCTGGGGGAGAGCGAGACTTCTCTCAAAARAAACAAAAA	2153
Db	4449	CACGCCACTGCATCTCTTTCTGGGTGACAGATGAGACTCCATCTCAAAAAACAAAAACA	4508
Qy	2154	CAAAAGAAATTAAGCAAATTAGACATTTCCAGAGAG	2187

```

Db      4509 AAACAAACAACAAAACAAACAAAAAACCACAG 4542
        ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 15
US-10-235-192A-34
; Sequence 34, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: Methods and Compositions for Identifying
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases
; TITLE OF INVENTION: and Disorders Associated Therewith
; FILE REFERENCE: WMI-011
; CURRENT APPLICATION NUMBER: US/10/235.192A
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-235-192A-34

Query Match      8.7%; Score 342.8; DB 13; Length 26928;
Best Local Similarity 57.5%; Pred. No. 4.9e-80;
Matches 848; Conservative 0; Mismatches 552; Indels 74; Gaps 10;

QY       735 TATTTTTTTTCTTGAAACAGGGTATCACTCTGTGTCGCCAGGCTGGAGTACAGTGGCGTAA 794
DB       3122 TTTTTTTTTTGGGAGTCAAGAGTCTTGCTCTGTGTCGCCAGGCTGGAGTGAGTGACACGA 3181

QY       795 TAATAGCTCACTGCAGGCTCC-CCTCCTGGGCTCAAGCAATCGCTGGCCTCAGCATCCT 853
DB       3182 TCTCAGCTCACTGCAGGCTCGGCTCCGGGTCAAGCTATTCTCTGCTCAGCCTCCC 3241

QY       854 GAGTAGCTGGGACTACAGGCTGTGTCACAGGCCACGACTTAAGTTTTTAAAAATGATTTT 913
DB       3242 AATAGCTGAGACTATAGGCACGCACATCCATGCCAGCTAATTTT-----TTATTTT 3295

QY       914 TGGTATAGAGAGGCTCTGTGATGTGTCTCAGGCTGATTTTATTGTTTGAGACAAGTGC 973
DB       3296 TAGTAGAGACGAGGTGTCTCCATGTGTGCCAGGTGTGCTTGAAC----- 3341

QY       974 TCATATGTTGCCATGATCCCCCACCTCCACTTCCCAAAGTGCTCATCTTATCTGTTC 1033
DB       3342 --CCTGCTCAAGTGATCCACCTGCCCTCGGCCTCCCAAAGTGTGGGATTCAGGCATG 3399

QY       1034 TTAGTCAGTTGACAGACATTTAGGTGTGTTTCCACTTTTTCACCATTTATGAATAATCTCC 1093
DB       3400 AGACACGCGCCGGGC-----TGCCTTGTCCCTTCTTAAATAGATTGTCATTGTAA 3454

QY       1094 AGTGAATATTCATGTATACATTTGTGTGGGCATATGTTTTTCATTTCTGTGTGGTTTATAT 1153
DB       3455 GCTGCTGATTTCTTTGGGACATTTGCTCCGTAACACTTTTCATAAAGCATCAGTGATTTCA 3514

QY       1154 CTAGGAGTGGAATTGCTGGATCCGGGTATATTTTGACAGGCAGAGTTTCAGGGGAAGAA 1213
DB       3515 CCATTTCTTCCACCCAGCTTCACCGTAATTTGTGTTGTTCTTCTGCTTCAATTCAGCA 3574

QY       1214 AAACTTTGGGAAAATGAAGCATGTTTAGAATCAGCAAGAGTGCAGGGGTTTTTCGGAGTT 1273
DB       3575 GAATTCATTTAGCTCTGATTAAGGCTCGCTTCAAACCTGATGCTTATCTTCTTAGTGCC 3634

QY       1274 TTAATTTATATTTCTGTTGACAAATGTGC-----AGTTTGATGAAGATACAGTTTACTTA 1328
DB       3635 TCAAACTACATCTCTGTTTCACTCATGTTATAGCAAGTTAGTGTGAGTTTATTTTGGTGCAC 3694

QY       1329 AGTCAGNAGTGAGAAATTAAGGCTTGAATAGGGCGTTTCAGAGTAAATCATGAAGCACTTT 1388
DB       3695 AAAAAATTTTTTAAATCCATGCGAGTCTTTTTTCATTAATACGCCATTTTCCATGAACATTTTC 3754

QY       1389 GAATACCAAAATTAAGGAGCTTGGCTGTAAACAAAATAATAAAAAATCAAAATTTTTTTTT 1448

```

```

Db      3755 GAAGACCCCTGTAGATGCTGTGTGTTTAAACACACCAGTTTA-----CAGTAATTTTTTT 3810
QY      1449 TTTTITTTTGGAGAAAGAGTCTTGTCTCTTCACTCCCTGGCTGGAGGGCAGTGTGTGATCTCA 1508
Db      3811 TTTTITTTTGGAGTAAAGTCTTGTCTCTGCGCCAGGCTGGAGTGCATTTGGCACACTCTCG 3870
QY      1509 GCTCACTGCAACTTTTGGCTCCCGGTTCAAGCAATTCCTGCTTTCAGCTTCCCAAGTA 1568
Db      3871 GCTCACTGCAACCTCTGCTCCCTGCTGGGTTCAAGCAATTTTCTGTCTCAGTCTCCGAGTA 3930
QY      1569 GCTGGGACTACAGGACATTCACACATGCCAGCTGATTTTGTATTTTGTAGAGATG 1628
Db      3931 GCTGGGATTTACAGGTTGTGCGCACCATGCTAGTAATTTATGTGTGTGTGTAGAGAGC 3990
QY      1629 GGATTTCACTTTGTGGCCCAAGCTGCTCAAACTTTTGTGCTCATATTTGTGTAAT 1688
Db      3991 GGGTTTCACTATTTGGCTAGGCTGGTCTCGAATCTCTCACCTTGTGATCGCGCCCTC 4050
QY      1689 ATTGTTCCITTTTGTAGGTAGGGCCCCCAGACCAAAATAAATAATCTTAGAATCCAAA 1748
Db      4051 GGCCTCCCAAGATTTGGGATTACAGGCGTGA-----GACTCTTGCACCTGGCC 4099
QY      1749 TCAGTGTGTTGTTGTTGACACCTGTCACTTGAGAACACAGTGTGACCGGCTCAGGAG 1808
Db      4100 TACAGTAATTTTATAGCAGCCTAGGCTAAGATAGGCCATTTCTGGGTATAAGAAATGTCATA 4159
QY      1809 TAGAGTGATCTCTGCTGAAAGAGAAATAGAAATGAAATATTTCTCGGGCCAGCGGTGG 1868
Db      4160 TACTGAACAGGCTCGCACTGTGATTAAGTCTGCAAGAGGCC-----GGGCAG 4210
QY      1869 TGGCTCATGCTGTGTAATCCAGCACTTTGGGGAGGCGAAGGCATGTGGATCACTGAGGTC 1928
Db      4211 TGGCTCATACCTGTAAATCCAGCACTTTGGGGAGGCGGAGGAGGTGGATCACTGAGGTC 4270
QY      1929 AGGAGTTCAAAACCCAGCCTGGCCACATGGTGAACCCCGTCTCTACTAAATAATCAAAA 1988
Db      4271 AGCAGTTCGAGACCCAGCCTGACCAATGGTGAACCCCGTCTCTACTAAATAATCAAAA 4328
QY      1989 AATTAGCTAAGTGTGTGGCGCATGCTGTAATCCC-----AGCTACTTG 2033
Db      4329 AATTAGCTGGGCGTGTGTAGTGCATGCTGTATCCCTAGCATGCACTTGGGAGCTACTTG 4388
QY      2034 GGAGGTCAGGCGAGAGAAATTTCTTGAACCCGGGAGGCGAGGTTGCGAGTGAAGCGAGAT 2093
Db      4389 GGAGGTCAGGCGAGAGAAATCACTTGTACTCAGGAGGCGGAGGTTGCGAGTGAAGCTGAGAT 4448
QY      2094 CACACCACTGCACTCCAGCCTGGGGAGAGAGCGAGACTTCCTCTCAAAAAACAAAAA 2153
Db      4449 CACGCCACTGCACTCTCTTCTGGGTGACAGAGTGAAGTCCATCTCAAAAAACAAAAA 4508
QY      2154 CAAAGAAATTAAGCAATTTAGACATTTGCAGAGAG 2187
Db      4509 AAACAAAAACAAAAACAAAAACCAACAG 4542

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:32:19 ; Search time 14547.2 Seconds
(without alignments)
8124.892 Million cell updates/sec

Title: US-10-016-725-15_COPY_1_3958

Perfect score? 3958

Sequence: 1 aggaatcattccattaaat.....gggacgtccacgcgcgg 3958

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	649.4	16.4	1238	12	BM563325 AGENCOURT
2	491.4	12.4	579	12	BI559696 603252566
3	450.2	11.4	471	28	AQ760013 HS_3202_A
4	334	8.4	754	12	BI627092 603075763

C	5	294.2	7.4	1910	11	BC035771	Homo sapi
	6	290.6	7.3	2076	11	BC035510	Homo sapi
C	7	290.6	7.3	2772	11	BC038630	Homo sapi
	8	283.4	7.2	2772	11	BC038630	Homo sapi
C	9	273.4	6.9	2076	11	BC035510	Homo sapi
	10	270	6.8	749	13	BU175345	AGENCOURT
	11	269.4	6.8	1172	12	BM546583	AGENCOURT
	12	269.4	6.8	1910	11	BC035771	Homo sapi
	13	259.4	6.5	929	9	AL580585	AL580585
	14	259.2	6.5	693	13	BU616025	BU616025
	15	255.8	6.5	2971	28	AF101960	AF101960
C	16	255.2	6.4	958	9	AL578531	AL578531
	17	255	6.4	694	9	AL596692	AL596692
	18	254.4	6.4	764	13	BU617735	BU617735
	19	252.4	6.4	940	13	EX424350	EX424350
	20	251	6.3	469	9	AL701463	AL701463
	21	251	6.3	560	9	AL708245	AL708245
	22	250.2	6.3	683	12	BI602109	BI602109
C	23	249.4	6.3	524	10	AW973992	AW973992
	24	249.2	6.3	580	12	BQ053726	BQ053726
	25	249	6.3	978	13	EX325356	EX325356
	26	248.4	6.3	1201	13	EX385602	EX385602
C	27	248.2	6.3	731	9	AI687343	AI687343
	28	247.2	6.2	775	12	BM905333	BM905333
C	29	246.8	6.2	521	13	BU617227	BU617227
	30	246.8	6.2	598	12	BQ017808	BQ017808
C	31	246.8	6.2	616	13	BU617236	BU617236
C	32	246.2	6.2	508	28	AQ418545	AQ418545
	33	246.2	6.2	958	9	AL578531	AL578531
C	34	245.6	6.2	969	13	EX415714	EX415714
C	35	245.6	6.2	2284	11	BC036238	BC036238
	36	245.4	6.2	467	28	AQ388230	AQ388230
C	37	245.4	6.2	785	12	BM721013	BM721013
	38	244.8	6.2	663	29	AG174449	AG174449
C	39	244.6	6.2	750	12	BM678642	BM678642
C	40	244.2	6.2	1201	13	EX385602	EX385602
	41	244	6.2	548	14	CB144391	CB144391
	42	243.8	6.2	755	28	BZ599689	BZ599689
C	43	243.8	6.2	2971	28	AF101960	AF101960
	44	243	6.1	826	28	BZ611620	BZ611620
C	45	243	6.1	929	9	AL580585	AL580585

ALIGNMENTS

RESULT 1
BM563325
LOCUS
DEFINITION
AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1238)
NIH-MGC <http://mgc.mci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LLAM12764 row: h column: 17
High quality sequence stop: 622.

BM563325 1238 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6564864 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5743936
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

```
FEATURES
source      Location/Qualifiers
1..1238
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5743936"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (Scorv site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match      16.4%; Score 649.4; DB 12; Length 1238;
Best Local Similarity 98.7%; Pred. No. 2.9e-90;
Matches 686; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 3266 TGGGGTGGTCCACGTCTCCCTGCC-CCGCGCTGGGGGACCGGGAAGACGGGACGGAGA 78
Db 20 TGGGGTGGTCCACGTCTCCCTGCC-CCGCGCTGGGGGACCGGGAAGACGGGACGGAGA 78

QY 3326 TGTAGTGGTGGCGGCCCCCGAGGTTACCACTGTTCTCAGAACTTCCCGAGTGC 3385
Db 79 TGTAGTGGTGGCGGCCCCCGAGGTTACCACTGTTCTCAGAACTTCCCGAGTGC 138

QY 3386 CCACCCACCGGTTCTCCGTGTGCCGAGGCGCGTCTCGGCTAGGCTCCGCGCCCGCAGC 3445
Db 139 CCACCCACCGGTTCTCCGTGTGCCGAGGCGCGTCTCGGCTAGGCTCCGCGCCCGCAGC 198

QY 3446 CCAAAACCGGGTCCCGAGCCCTTCAGAGAGAAAGCTCCGAGCGGGATCCGGGCGAG 3505
Db 199 CCAAAACCGGGTCCCGAGCCCTTCAGAGAGAAAGCTCCGAGCGGGATCCGGGCGAG 258

QY 3506 AGGCCAGCGGGGTGGAGAGAGAGCTGAGAGGAGAAACAGAGGGGAGGGGAGCGAG 3565
Db 259 AGGCCAGCGGGGTGGAGAGAGAGCTGAGAGGAGAAACAGAGGGGAGGGGAGCGAG 318

QY 3566 GAGCTGGCGGCGAGAGGAAACAGCAGATTGCGCGAGCCCAATGGCAACGGCAGACGAGGT 3625
Db 319 GAGCTGGCGGCGAGAGGAAACAGCAGATTGCGCGAGCCCAATGGCAACGGCAGACGAGGT 378

QY 3626 GGCACCAAAATTCCTTCGGGCAATGACGAGCGGAGTTTACAGAGCCTCATTAGCATTT 3685
Db 379 GGCACCAAAATTCCTTCGGGCAATGACGAGCGGAGTTTACAGAGCCTCATTAGCATTT 438

QY 3686 CCCCAGAGCAGGGGACGAGGCGCGGTGGTGTGCTGTGCTGTCTCTCT-CCGCCCTCCGCGCCT 3745
Db 439 CCCCAGAGCAGGGGACGAGGCGCGGTGGTGTGCTGTGCTGTCTCTCTCTCTCCGCCCTCCGCGCCT 498

QY 3746 CCGGCGCCTGTCTGCGGTCCGCGAGGCTTCGGGCTCTGTCTCTCT-CCGCCCTCCGCGCCT 3804
Db 499 CCGGCGCCTGTCTGCGGTCCGCGAGGCTTCGGGCTCTGTCTCTCTCTCTCCGCCCTCCGCGCCT 558

QY 3805 TACCTCCAGCGGGACCGCGCGGCGAGTCACTCTCTGCACTTTGCGCCCTCTGTTGGGAG 3864
Db 559 TACCTCCAGCGGGACCGCGCGGCGAGTCACTCTCTGCACTTTGCGCCCTCTGTTGGGAG 618

QY 3865 CGGATAAAAGGGGCTGAGGAAATACCGGACAGGTCACCGCTTGGCAGCTCTAGCCTTT 3924
Db 619 CGGATAAAAGGGGCTGAGGAAATACCGGACAGGTCACCGCTTGGCAGCTCTAGCCTTT 678

QY 3925 AAATTTCC-GGCTCGGGGACCTTCCAGCAGCGCGG 3958
Db 679 TAATTTCCCGGGCTCGGGGACCTTCCCGCACCAGCGG 713
```

```
RESULT 2
BI559696      579 bp      mRNA      linear      EST 05-SEP-2001
603252566F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5294921 5',
mRNA sequence.
BI559696
BI559696.1 GI:15447010
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 579)
NIH-MGC http://mgs.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11746 row: c column: 18
High quality sequence stop: 576.
Location/Qualifiers
1..579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5294921"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(GTCGAG); Oligo-dT primed using primer
5'-TTTTTATTTTATTTTATTTT-3', size-selected for average
insert size 2.2 kb and normalized to 50x. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match      12.4%; Score 491.4; DB 12; Length 579;
Best Local Similarity 97.6%; Pred. No. 8.4e-66;
Matches 562; Conservative 0; Mismatches 7; Indels 7; Gaps 6;

QY 3229 GGAGCGGGGCTTCCCGGTCCCAAGCTCCAGATCTCGGGTGGCTGCCAGTCTCCCT 3288
Db 4 GGAGCGGGGCTTCCCGGTCCCAAGCTCCAGATCTCGGGTGGCTGCCAGTCTCCCT 63

QY 3289 GCCACGCGCTGGGGGACGGGAACACGGGACGGAGATGTTAGT-GGTGGGGCCCCCG 3347
Db 64 GCCACGCGCTGGGGGACGGGAACACGGGACGGAGATGTTAGTGGGTGGGGCCCCCG 123

QY 3348 AGGG--TTCAACAATGTTTCTTGAGAAACTTTCCCAAGTCCCAACCCCGTCTCCG 3405
Db 124 AGGGTTTCAACAATGTTTCTTGAGAAACTTTCCCAAGTCCCAACCCCGTCTCCG 183

QY 3406 TGCCGAGGGCGGTCTCGGGTAGGCTCCGGGCCCGGACCCCAACCGGGTCCCGAGCC 3465
Db 184 TGCCGAGGGCGGTCTCGGGTAGGCTCCGGGCCCGGACCCCAACCGGGTCCCGAGCC 243

QY 3466 CTTTCCAGAGAGAAAGCTTCCCGAGCGGGATGCGGGGACAGAGCCAGCGGGTGGAA 3525
Db 244 CTTTCCAGAGAGAAAGCTTCCCGAGCGGGATGCGGGGACAGAGCCAGCGGGTGGAA 303

QY 3526 GAGAAGCTGAGAGGAGAAACAGAGGGGAGGG-GGAGCGAGGAGCTGCGCGGACAGGGAA 3584
```


full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

```

ORIGIN
Query Match      8.4%; Score 334; DB 12; Length 754;
Best Local Similarity 99.7%; Pred. No. 1.1e-41;
Matches 345; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3613 GGCAGGACGAGTGGCCAAATTCCTTCGCGCAATGACGAGCGGAGTTTACAGAAGC 3672
Db 1 GGCAGGACGAGTGGCCAAATTCCTTCGCGCAATGACGAGCGGAGTTTACAGAAGC 60

Qy 3673 CTCATTAGCATTTCCAGAGGCGAGGCGGAGGCGAGGCGCGGTTGTGTGCGT 3732
Db 61 CTCATTAGCATTTCCAGAGGCGA-GGCGAGGCGAGGCGCGGTTGTGTGCGT 119

Qy 3733 GTCGCGAGCATCCCGCGCGCCCTGCTCGGFGTCGCGGAGCTCGGCTCTGTCTCTCC 3792
Db 120 GTCGCGAGCATCCCGCGCGCCCTGCTCGGFGTCGCGGAGCTCGGCTCTGTCTCTCC 179

Qy 3793 CCCTCCCGCCTTACCTTCACGCGGAGCGCGCGCGGAGTCAACTCTCTCGCACTTGGC 3852
Db 180 CCCTCCCGCCTTACCTTCACGCGGAGCGCGCGCGGAGTCAACTCTCTCGCACTTGGC 239

Qy 3853 CTGCTTGGCAGCGGATTAAGGGGGCTGAGGAATACCGACGCTACCGTTGCCA 3912
Db 240 CTGCTTGGCAGCGGATTAAGGGGGCTGAGGAATACCGACGCTACCGTTGCCA 299

Qy 3913 GCTCTAGCCTTTAAATTCGCGGTCGCGGACCTCCACGCGACCGGG 3958
Db 300 GCTCTAGCCTTTAAATTCGCGGTCGCGGACCTCCACGCGACCGGG 345

```

```

RESULT 5
BC035771/c
LOCUS
DEFINITION
Homo sapiens RAD1 homolog (S. pombe), mRNA (cDNA clone
IMAGE:5582524), with apparent retained intron.
ACCESSION
BC035771
VERSION
BC035771.1 GI:23958889
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1910)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22398257
12477932
REFERENCE
2 (bases 1 to 1910)
Strausberg, R.
AUTHORS
Direct Submission
TITLE

```

JOURNAL

REMARK
COMMENT

Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-x@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tauregeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate; 79 Row; d Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 this clone has the following problem: retained intron.

FEATURES

Location/Qualifiers

1..1910

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5582524"
 /tissue_type="Testis, embryonal carcinoma"
 /clone_lib="NIH MGC_92"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

```

Query Match      7.4%; Score 294.2; DB 11; Length 1910;
Best Local Similarity 64.5%; Pred. No. 8.2e-36;
Matches 496; Conservative 0; Mismatches 253; Indels 20; Gaps 3;

Qy 1414 TGTAAACAAATAATAAATAATCAAAATTTTTTTTTTTTGTGAGAAAGAGTCTTGCTC 1473
Db 1176 TGCAGCTAAACAAATAATTTTTTTTAAATTTTTTTTGTGAGCGAGTCTTGCTC 1117

Qy 1474 TTTACCCCTGGCTGGAGGCGACTGGTGTGATCTCAGCTCAGTCAGCACTTCGCGCTCCCGG 1533
Db 1116 TGTACCCAGGCTGGAGTGCATGTGCAATGTGCAATCTCGGCTCAGTCAACCCCGCTTCAG 1057

Qy 1534 GTTCAAGCAATTCCTCCTGCTCAGCCTCCCAAGTAGCTGGGACTACAGCACTTCCACCC 1593
Db 1056 TTTCAGTGATTTCTCTGCTCAGCCTCTCAAGTAGCTGGGATTACAGACACCCACACC 997

Qy 1594 ATGCCCGAGCTGATTTTCTATTATTTTAGTAGAGAT-GGATTTTCATCTTTTGGCCAGCT 1652
Db 996 ACACCCAGCTAATTTTGTATTTTAGTAGAGGTAGGGGTTTTCACCATGTTGGCCAGGCT 937

Qy 1653 GGTCTCAAACTTTTGTCTGTCTAATTTTGTAACTATTGTTCTTTTGTGCTGAGTAGGG 1712
Db 936 GATCTCGAATCTCTGACCTCAGA-----TGATCCCTGCTCTGCGC 895

Qy 1713 CCCCGACACCAAAATAAATAATCTTAGAATCCAAATCAGTGTGTGTGTTGACCACTGT 1772
Db 894 TTCCCAAAATGCTGGGATTACAGGTGTGAGCCACCACTCCCGGCCCTAAAAATAATGTTT 835

Qy 1773 CACTGTGAGAACCAACAGTGTGACCGGGCTCAGGAGTAGAGGTGATCTCTGCTCGAAGA 1832

```

Db 834 AAATGTACAAAATAATCCATCTTACTAGCTGATTTATATTTTGTATTAATCTTTTAA 775
 QY 1833 GAAATAGAAATG-AAAATATTTCTCGGGCCAGGGCGTGGTCTCATCGCTGTAATCCAGC 1891
 Db 774 GATTCAACTTTTAAATAATATTAGCCAGGCATGATAGCTACACCCCTGTATCTCCAGC 715
 QY 1892 ACTTTGGAGGCCCAAGCATGTGGATCACCTGAGGTCAGGAGTTCAAAACCCAGCTGGCC 1951
 Db 714 CTTTTGGAGGCCGAGCGGGTGGATTGCTTGAGCTCAGGAGTTCAGATTAGCTTAGGG 655
 QY 1952 AACATGCTGAAACCCCGTCTCTACTAAATAACAAAATAGCTTAAGTGTGGTGGCGCA 2011
 Db 654 ACATATAGTAAACCCCATCTCTACAAAATAATACAAAATATTAGCAGGCGATGAGGCATG 595
 QY 2012 TGCTCTTAATCCAGCTACTTGGGAGGGTGAGCGAGGAATTTCTTGAACCCGGGAGGC 2071
 Db 594 CACCTGTAGCTAGCTATTGGGAGCTGAGATGGGAGGATGCTTGGAGCTTAGGAGAT 535
 QY 2072 AGAGGTGAGTGAAGCGAGATCACACCTGCTACCTCAGCTGCGGGGAGAGAGGAGAC 2131
 Db 534 TGAGGCTGAGTGAAGTGAATCTTACAACTGTACTTCCAGCTGGGCAACAGAGCAAGAT 475
 QY 2132 TTCTCTTCAAAAACAAAACAAAAGATTAAGCAATTTAGACATTG 2180
 Db 474 CTGTCTCCAAAATAAAAATAAAAACTGAGATCAAAAATAAATTAG 426

RESULT 6
 BC035510
 LOCUS
 DEFINITION Homo sapiens sialophorin (gpl15, leukostialin, CD43), mRNA (cdna clone IMAGE:5094058), with apparent retained intron.
 ACCESSION BC035510
 VERSION BC035510.1 GI:23271054
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2076)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, I., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, X.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 MEDLINE
 PUBMED
 REFERENCE 2 (bases 1 to 2076)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, P.J., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.D., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: P Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5094058"
 /tissue_type="Lymph, Burkitt lymphoma"
 /clone_lib="NIH MGC 8"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 7.3%; Score 250.6; DB 11; Length 2076;
 Best Local Similarity 67.2%; Pred. No. 2.8e-35;
 Matches 498; Conservative 0; Mismatches 214; Indels 29; Gaps 5;
 QY 1421 AAAATAATAAAAAATACAAATTTTTTTTTTTTGGAGAAAGAGCTTGTCTTTTACC 1480
 Db 741 ACACACGTGAATAGATGCTCTTTTTTTTGGAGATGGAGTTTTCACCTCTGTACC 800
 QY 1481 CTGGCTGAGGGGAGTGGTGTGATCTAGCTACCTGCACTTCCCTCCCGGTTCAAG 1540
 Db 801 CAGGCTGAGTGCATGGCGTGTCTTGGCTCACCAGAACCTCCGCTCTCTGGTTCAA 860
 QY 1541 CAATTTCTCTCTTCCAGCTCCCAAGTAGCTGGGACTACAGCAGCTCCACCATGCCCA 1600
 Db 861 CAATTTCTCTCTTCCAGCTCCCAAGTAGCTGGGACTACAGCAGCTCCACCATGCCCA 920
 QY 1601 GCTGATTTTGTATTTTGTAGTGGGATTTTCACTTTTGTGGCCAGCTGGTCTCAA 1660
 Db 921 GAT-AAATTTTGTATTTTGTAGTGGGATTTTCACTTTTGTGGCTGGTCTTGA 979
 QY 1661 ACTTTTGTCTGTATTTTGTGTAACTATTTTCTTTTGTGGTGGGCCCCCAGA 1720
 Db 980 ACT-----CCGAGCTCTGTGATCCGCTATGCTCGGCTCCCAAGTGTGGGATT 1031
 QY 1721 CCAAAAAAATAAATCTTAGAATCAAAATCAGTGTGTGTGGTTGACCACTGTCACTGAG 1780
 Db 1032 ACATGCTGTAGCCACCCTGCGCCGACATAGATGCTTTTAAATTTTCTGGAGAAAGCA 1091
 QY 1781 AACACAGTGTGACCGAGGCTCAGGAGTAGAGTGATCTCTGTCGAAAGAGAAATAGA 1840
 Db 1092 AAGCAAAAGAGCAGTGGATATTTTAA-----ACTAAAAAGGAAACAAA 1137
 QY 1841 ATGAAATATTTCTCCGGCCAGGGCTGGTGTCTC-ATGCCGTGTAATCCACACTTTGGG 1899
 Db 1138 AAAGGAGATAGACGAGCCAGAGCTGGTGTCTCAACGTCTGTATCCACACTTTGGG 1197
 QY 1900 AGGCCAAGGCATGTGGATCACCTGAGGTTCAGGAGTTCAAAACACGCTGGCCACATGGT 1959

Db	1198	
----	------	--

Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Amuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 73 Row: 1 Column: 4
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..2772
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5263792"
 /tissue type="Brain, hippocampus"
 /clone_lib="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 7.2%; Score 283.4; DB 11; Length 2772;
 Best Local Similarity 66.1%; Pred. No. 3e-34;
 Matches 486; Conservative 0; Mismatches 216; Indels 33; Gaps 4;
 QY 1441 TTTTCTTTTCTTTTGTGAGAAAGTCTTGTCTTCTTCCCTGGCTGGAGGCGAGTGGTG 1500
 DB 2071 TTTCTTTTCTTTTCTGAGATGGAGTCTTCTCTGTCTGTCGCCAGCT-GAGTGCATGGTG 2129
 QY 1501 TGATCTCAGCTCACTGCAACTTTCCTCCCGGTTCAGCAATTCCTGCTTCAGCCT 1560
 DB 2130 TGATCTCGGCTCACTGCAACTTTCCTCCCGGTTCAGCAATTCCTGCTTCAGCCT 2189
 QY 1561 CCCAAGTAGCTGGAGCTACAGGCACTTCCCACCATGCCCCAGACCAAAATAATCTTAG 1620
 DB 2190 CCGGAGTAGCTGGGTCTACAGGCACTGCACTAGCCTGGCTGCTGATTTTGTATTTAG 2249
 QY 1621 TAGAGATGGGATTTCACTTTGTGGCCAACTGGTCTCAAACTTTTGTCTCATATTTG 1680
 DB 2250 TAGAGGCAAGTTTCCACATGTGTGTCAGGCTGGTTGAACTCTCTGACCTCAGTGA 2309
 QY 1681 TTGTAACTATTTCTCTTTCTGAGTAGGGCCCCAGACCAAAATAATCTTAG 1740
 DB 2310 ACCGGCTCGGCTCTCAAGGTGTGTGATTAAGGATGAGCCACCTCTCGGCATAT 2369
 QY 1741 AATCCAAATCAGTGTGTGTGTTTGTGACCACTGTCTTGAACCAACAGAGTGTGACAGGGC 1800
 DB 2370 AACACTTTTCAGTAAAT-----CATCTCTATTATTATTTAGCAGAGTTCACCTCTA 2422
 QY 1801 CTCAGAGTAGAGGTGATCTCTCTCGAAGAGAAATAGAAATATTCTTCGGGGCC 1860
 DB 2423 ACAAGCAAGATAGTTGATATATCAAGAA-----TCCTAGCT 2458
 QY 1861 AGCGGTGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCAATGTGGATCAC 1920
 DB 2459 GGGAGCTGGCTCATGCTGTAATCCACCACTTTGGGAGGCTGAGCGAGTGGATCAC 2518
 QY 1921 CTGAGTCAAGGAGTTCAAAACCAAGCTGGCCAACTGGTGAACCCCGCTCTTACTAAAA 1980
 DB 2519 CTGAGTCAAGGAGTTGAGACCAAGCTGGCCAACTGGTGAACCCCAATATATCTAAAT 2578
 QY 1981 ATACAAAAATAGCTAAGTGTGGTGGCGATGCTGTAAATCCAGCTACTTGGGAGGCT 2040
 DB 2579 ATAC-AGAAATTTGGCTGGGCGTGTGGCGGCCCACTGTGCTCCGCTACTTGGGAGGCT 2637
 QY 2041 GAGGCAAGGAGATTTCTTCAACCCGGAGGCGAGAGGTTCCAGTGAAGCAGATCACACCA 2100
 |||||

Db 2638 GAGGCAGGAGATCTGTTCAACCCGGAGGTGAGTTCGAGCGAGCCGAGATGGCACCA 2697
 QY 2101 CTGACTCCAGCTCGGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAAAGA 2160
 Db 2698 TTGCACTCCAGCTCGGAGACAGAGCAAACTCGTCTTACCAAAAAAAGAAAGA 2757
 QY 2161 ATTAGCAAAATAGA 2175
 Db 2758 AAAAAAAGAAAAA 2772

RESULT 9

BC035510/c

LOCUS

DEFINITION

Homo sapiens sialophorin (gpl15, leukosialin, CD43), mRNA (cdna

clone IMAGE:5094058), with apparent retained intron.

ACCESSION BC035510.1 GI:23271054

VERSION BC035510

KEYWORDS HTC

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 2076)

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,

Carninci, P., Frange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,

Armonon, K.D., Mullihy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smal, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2076)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisec.nih.gov/>Contact: nisc_mgc@nih.gov

Akter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,

Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: p Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES
 source
 1. 2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5094058"
 /tissue_type="Lymph. Burkitt lymphoma"
 /clone_lib="NIH_MGC_8"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 6.9%; Score 273.4; DB 11; Length 2076;
 Best Local Similarity 67.4%; Pred. No. 1.2e-32;
 Matches 487; Conservative 0; Mismatches 206; Indels 30; Gaps 6;
 QY 1444 TTTT...TTCAGAAAGAGCTTCTCTT-TCACCTGCTGGAGGCGAGTGTGTG 1502
 Db 1446 TTTT...TTCAGATGGAGTTTACTTTGTCGCCAGGCTGGAGTGCATGGCGTG 1387
 QY 1503 ATCTCAGCTCACTGCACTTTCGCTCCGGGTTCAAGCAATCTCTGCTCAGCCTCC 1562
 Db 1386 ATCTCGGCTCACTGCAACCTCCACTCTCTGGGTTCAAGGGATTCTCTGCTCAGCCTCC 1327
 QY 1563 CAAGTAGCTGGGACTACAGGCACCTCCACATGCCCCAGCTGATTTTGTATTTTAGTA 1622
 Db 1326 CAAGTAGCTGGATCACAGTGCCTGCCACACAGCCAGCTAATTTTGTATTTT 1271
 QY 1623 GAGATGGATTTCACTTTGTGGCCAGCTGGTCTCAAACTTTTGTGTCATAAATGTT 1682
 Db 1270 GAAACAGGGTTTCCACATGTTGTGTCAGCTGGTCTTGAACCT-----CCTGACCTCAG 1219
 QY 1683 GTAACATATTGTTCTTTTGTGAGTAGGAGGCCCCAGACCAAAAAAATAAATCTTAGAA 1742
 Db 1218 GTGATCCACTGCTCGGCTCCCAAGTGTCTGGGATTACAGACGTTGAGCCACCAACGCTC 1159
 QY 1743 TCCAAATCAGTGTGTGTTTGTGACCACTGTCAGTTCAGAACCAACAGTGTGACAGGGCT 1802
 Db 1158 TGCCCTGCTCTATCTCTCTTTTGTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTT 1099
 QY 1803 CAGGAGTAGAGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCGGGCCAG 1862
 Db 1098 TTTGCTTTGCTTTTCTCTCC-----AGAAATTAAGACATCTATTGTCGG 1052
 QY 1863 GCGTGTGCTCATGCTCTAATCCAGCACATTTGGGAGCCAGGCGATGTGATCACT 1922
 Db 1051 GCACGTGTGCTCAGCATGTATCCAGCACATTTGGGAGCGGAGGATACGATCAGCA 992
 QY 1923 GAGGTGAGGAGTTCAAAACAGCTGGCCAAATGTTGAAACCCGCTCTCTACTAAAT 1982
 Db 991 GAGGTGCGGAGTTCAAGACAGCTGACCAATGTTGAAACCCGCTCTCTACTAAAT 932
 QY 1983 ACAAAATATTAGTGTGTTGCGCATGCTGTAATCCAGCTACTTGGGAGGTTGA 2042
 Db 931 AC--AAATATTCCGGGTTTGTGTTGTCATGCTGTAATCCAGCTACTTGGGAGGTTGA 874
 QY 2043 GCAGGAGATTTCTTGACCGGAGGCGAGGTTGAGTGAAGGAGATCACACCT 2102
 Db 873 GCAGGAGATTTGTTTGAACCGGAGGCGAGGTTTCTGTTGAGCAAGATCACGCCATT 814
 QY 2103 GCATCCAGCTGGGGAGA--GAGCGAGACTTCTCTCAAAAAACAAAAAATAAAGA 2160
 Db 813 GCATCCAGCTGGGTTAACAAGTGAATCTCTCAAAAAAATAAAGACATCT 754
 QY 2161 ATT 2163

Db 753 ATT 751

RESULT 10
 BUL75345

LOCUS
 DEFINITION
 BUL75345
 5', mRNA sequence.
 BUL75345
 BUL75345.1. GI:22689329
 EST.
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

NIH-MGC <http://mgs.nci.nih.gov/>.
 (bases 1 to 749)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13202 row: i column: 22
 High quality sequence stop: 530.

FEATURES
 source

1. 749
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6011997"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: Sali; Cloned unidirectionally; oligo-dr primed. Average insert size 2.5 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.8%; Score 270; DB 13; Length 749;
 Best Local Similarity 64.5%; Pred. No. 7.4e-32;
 Matches 462; Conservative 0; Mismatches 230; Indels 24; Gaps 3;
 QY 1443 TTTT...TTCAGAAAGAGTCTTCTCTTCCCTGCTGGAGGCGAGTGTGTG 1502
 Db 52 TTTT...TTCAGAAAGAGTCTTCTCTTCCCTGCTGGAGGCGAGTGTGTG 111
 QY 1503 ATCTCAGCTCACTGCACTTTCGCTCCGGGTTCAAGCAATCTCTGCTTCAAGCTCC 1562
 Db 112 ATCTTGGCTCACTGCAACCTCTGCTCTCTGCTTCCGATTTCCCATCTTAGCTCC 171
 QY 1563 CAAGTAGCTGGGACTACAGGCACCTTCCACCATGCCCCAGCTGATTTTGTATTTTAGTA 1622
 Db 172 TCAGTAGCTGGGACACAGGCGTGTGCCACACATCCAGCTAATTTTGTATTTTGTG 231
 QY 1623 GAGATGGGATTTCACTTTTGTGGCCAGCTGCTCAAACTTTTGTCTGTCTATAATTGT 1682
 Db 232 GAGACGGGTTTCACTATGTTTGGCAGGCTGTGTGCTGACTCTCTGAC----- 278
 QY 1683 GTAACATTTGTTCTTTTGTGAGTAGGCGGCCCGCCAGACCAAAAAATAAATCTTAGAA 1742
 Db 279 -----CTAAAGCGATCTGTCTGCTCGGCTCCCAAGTGTGGGATAAGTCTGGGAT 332
 QY 1743 TCCAAATCAGTGTGTGTTTGTGACCACTGTCACATGTGAGAACCAACAGTGTGACCGGCT 1802

Query Match	6.8%;	Score 269.4;	DB 12;	Length 1172;
Best Local Similarity	97.8%;	Pred. No. 7.1e-32;		
Matches 273; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;
3680 GCATTTCCCACAGCGCAGGGGCGAGGGGCAGAGGCCGGGTGGTGTGGTGTCGGTCTCGGCA	3739			
6 GGAATTCCTCGGGATGCAAGGGCGAGGGGCAGAGGCCGGGTGGTGTGGTGTCGGCA	65			
3740 GCATCCCGCGGCGCCTGCTGCGGTGCGCGCGAGCCTCGGCGCTCTGCTCCTCCCCCTCCC	3799			
66 GCATCCCGCGGCGCCTGCTGCGGTGCGCGCGAGCCTCGGCGCTCTGCTCCTCCCCCTCCC	125			
3800 GCCCTTAACCTCCACGCGGGAACGCGCGCGGCAGTCAACTCCCTCGCACATTGCCCCCTGCTT	3859			
126 GCCCTTAACCTCCACGCGGGAACGCGCGCGGCAGTCAACTCCCTCGCACATTGCCCCCTGCTT	185			
3860 GGCAAGCGGATAAAGGGGGCTTAGGAAATACGGACACGGTACC CGTTGCCAGCTCTAG	3919			
186 GGCAAGCGGATAAAGGGGGCTTAGGAAATACGGACACGGTACC CGTTGCCAGCTCTAG	245			
3920 CCTTTAAATTCCTCCGGCTCGGGGACCTCCACGACCGCGG	3958			
246 CCTTTAAATTCCTCCGGCTCGGGGACCTCCACGACCGCGG	254			

RESULT 12	BC035771	1910 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	BC035771				
DEFINITION	Homo sapiens RAD1 homolog (S. pombe), mRNA (cDNA clone IMAGE:5582524), with apparent retained intron.				
ACCESSION	BC035771				
VERSION	BC035771.1	GI:23958839			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1910)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,				

Scapellato, M., Soares, A.B., Bonardo, M.F., Casavola, A.L.,
Scheetz, R.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carrinuci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.-J.,
Abramson, R.D., Mullahy, S.J., Bossak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallegange, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 1910)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapsb-x@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc.

Email: cyapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.

Query Match	6.6%	Score 259.4;	DB 9;	Length 929;
Best Local Similarity	62.2%	Pred. No. 2.8e-30;		
Matches 507;	Conservative 14;	Mismatches 245;	Indels 49;	Gaps 6;

QY	1441	TTTTTTTTTTTTTTTGGAAGAGCTTGTCTCTTTCCCTCGCTGGAGGGCAGTGCGTG	1500
Db	1116	TTTTTTTTTTTTTTTGTGTCGAGTCTCTCCCTGTGCCAGGCTGAGCACAGTGGCG	175
QY	1501	TGATCTACGTCATGCAACTTTTCGCTCCCGGGTTCAAGCAAAATCTCTCTTCAGCCT	1560

Query Match	6.8%; Score 269.4; DB 11; Length 1910;
Best Local Similarity	64.8%; Pred. No. 5.3e-32;
Matches	471; Conservative 0; Mismatches 236; Indels 20; Gaps 4;
QY	1441 TTTTNTTTTTTTTTTTGAGNAGAGTCTTGCTCTTTTCAACCTGGCTGGAGGGGAGTGGTG 1500
Db	451 TTTTNTTTTTTTTTTTGAGNAGAGTCTTGCTCTTTTCAACCTGGCTGGAGGGGAGTGGTG 510
QY	1501 TGAATCTCAGCTCACTGCACCTTTTCGCTCCGGGTTCAAGCAATTTCTCTGCTTCAGCCT 1560
Db	511 GGATTCAGTTCACTCCAGCCTCAATCTCTTAGCTCAAGCAATCTCCCATCTCAGCTT 570
QY	1561 CCCAGTAGCTGGGACTACAGGCACTTCCCAACATGCCAGCTGATTTTTGTA - TTTTTA 1619
Db	571 CCCAAATAGCTACGACTACAGGTGCAATGCTCCATGCCCTGTAATTTTTGTATTTTTT 630
QY	1620 GTAGAGATGGGATTTTCACTTTTGTGGCCAACTGGTCTCAAACTTTTTCGCTGCTCAATTT 1679
Db	631 GTAGAGATGGGTTTTTACTATGTTCCCTAGGCTAATCTTGAAC - - - - TCCCTGAGCTCAAG 686
QY	1680 GTTGTAACATATTTGTCCTTTTGTGAGTGGGCCCCCAGACCAAAAAAATAAATCTTTA 1739
Db	687 CAATCCACC CGCCTCGGCCCTCCAAAGGGCTGGGATACAGCGCTGAGCTATCATGCCCTG 746
QY	1740 GAATCCAAATCAGTGTGTGTTTTGACCACTGCTACTGAGAACCAACAGTGTGACAGGG 1799
Db	747 GCTTAATGTATTTTTTAAAGTTGAACTTAAAGAAGAAATTTAAACCAAAATAAAATCAGCT 806
QY	1800 CCTCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGC 1859
Db	807 CAGTAAGGATGATATTTTGTACATTTAAAACATA - - - - - TTAATTTTAGGC 852
QY	1860 CAGCGGTGTGTGCTCATGCTGTAATCCAGCACTTTGGGAGGCCAAGGCATGTGGATCA 1919
Db	853 CGGAGTGTGTGCTCACACTGTAATCCAGCATTTTGGAGCCGAGGCAGGGGGATCA 912
QY	1920 CCTGAGGTCAGG - AGTTCAAACACAGCTGGCCCAACATGGTGAACCCCGCTCTTACTAA 1978
Db	913 TCTGAGGTCAGAAGTTCCAGATCAGCTGGCCCAACATGGTGAACCCCTACTCTACTA 972

```
Db 176 TGATVTCAGCTCGCTGCAACCTCTCCCTCCCGTGTTCAGCAATTTCTCTGCTCAGCCT 235
QY 1561 CCCAAGTAGCTGGGACTACAGGACTTCCACCATGCCAGCTGATTTTGTATTTTAG 1620
Db 236 CCCGAGTGCCTGGGACTACAGGCTGTGCCACACACTCAGCTAATTTTGTATTTCCAG 295
QY 1621 TAGCATGGGATTTTCACTTTGTGGCCAGCTGTCTCAAACTTTTGTCTGCATATAATTG 1680
Db 296 TAGAGACTG--TTTCACCATGTTGTGTGAGCTGGTCTTGAACCTCTGACTTCGTGA--- 349
QY 1681 TTGTAACTATTGTTCTCTTTGCTGAGGTAGGGCCCCAGACCAAAAATAAATCTTAG 1740
Db 350 -----TCCGCCACCTCAGCCTCCCAAGTGTGGGATTACAGCGCAG 392
QY 1741 AATCAAAATCAGTGTGTTGGTTTGACCACTGTCATTTGAGAACCACAGTGTGACCGGCG 1800
Db 393 CCACC-----TTGCCCGCGCCGAAGTTAAAGATGCGAGTTGGGATAAGACC 437
QY 1801 CTCGAGGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAAATATTTCTCCGGGCC 1860
Db 438 ATCATAGGATATTTCTTCCCTT-----TACTCATAAAGTTGTGACTTTGGCC 489
QY 1861 AGCGGTGTGGCTCATGCTGTAATCCAGACATTTGGAGGCCCAAGCATGTGTGATCAC 1920
Db 490 GGGCGGATGGCTCACGCTGTTATCTAGTACTTTGGAGGCTCAGCGGGCAGATCAC 549
QY 1921 CTGAGGTGAGGATTTCAAAACAGCTGGCCAAACATGTTGAAACCCCTCTCTACTAAA 1980
Db 550 CTGAGGTGAGGATTTGTACAGCTGGCCAAACATGAAACCCCTCTCTACTAAA 609
QY 1981 ATACAAAAATTAGCTAAGTGTGGGGCATGCTCTGTAATCCAGCTACTTTGGAGGGT 2040
Db 610 ATAC--AAACTAGCCAGTGTGTGTGGGCACCTGCGAGTCCAGCTATTCGGGAGCT 667
QY 2041 GAGGAGGAGAAATTTCTTGAACCCGGGAGGAGAGGTTGAGTGAAGCAGATCACCA 2100
Db 668 GAGGAGGAGAAATCACCCGAACTGGGAGGAGAGGTTGTAGCGAGCGAGATCTTGCCA 727
QY 2101 CTGCACTCCGCTGGGAGGAGAGCGAGACTCTCTCTCAAAAAACAAAAACAAAGA 2160
Db 728 TCACACTCCGCTGGAGGAGCAGAGCAAACTCTGTCCAAAAA--AAAAAAGT 786
QY 2161 ATTAGCAAAATTAGCATTTGACAGAGAACCTTGAAGGGGGTTCAGACCGTACAGATTTC 2220
Db 787 TGTAGCTTAAATAATCTTCWACGAGAGCTTTCTTTTGTCTCTAAACCTTCTCTTT 846
QY 2221 TGTGCCACATGCCAAGTACTCTGAGGCGATGACTG 2255
Db 847 TTTAAWCTGTGTGCGGRTTTTGAGGGGAAACAG 881
```

```
RESULT 14
BU616025
LOCUS
DEFINITION
UI-H-DF0-ben-m-20-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone
UI-H-DF0-ben-m-20-0-UI 3', mRNA sequence.
ACCESSION
BU616025
VERSION
BU616025.1 GI:23282240
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 693)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tissue Procurement: Dr. Jose Mercuende
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
```

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 9-305, >ALU (matched complement) 326-387, >L1MB7#LINE/L1
(matched complement) 388-676, >ALU
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers

```
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-ben-m-20-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAGCGTC.
TAG_TISSUE=Subchondral bone
TAG_LIB=UI-H-DF0
TAG_SEQ=GTTAAGCGTC"
```

ORIGIN

```
Query Match 6.5%; Score 259.2; DB 13; Length 693;
Best Local Similarity 64.7%; Pred. No. 3.6e-30;
Matches 465; Conservative 0; Mismatches 216; Indels 38; Gaps 4;
QY 1441 TTTTCTTTTGTGAGAAAGATCTTGTCTTTACCTGGCTGAGGGCAGTGGTG 1500
Db 1 TTTTCTTTTGTGAGCAGAGTTTGTCTATCACCAGCTGGAGTGCAGTGCA 60
QY 1501 TGATCTCAGCTCAGTCAACTTTTCCTCCCGGTTCAAGCAATTTCTCTGCTCAGCCT 1560
Db 61 CAATCACAGTTCACAGCAGCTGTAACTGTGGGCTCAAGCGTCTCCACCTTAGCCT 120
QY 1561 CCCAAGTAGCTGGGACTACAGGCACTTCCCACCATGCCAGCTGATTTT---GTATTTT 1617
Db 121 CCCGGTAACCTGGAACCTGCAGGTGTGCCACTATGCCAGCTAATTTTAAAAAAT 180
QY 1618 TAGTAGATGGGATTTTCACTTTGTGGCCAGCTGTTCTCAAACTTTTGTGTCATAA 1677
Db 181 TTATAGATAGAGTCTCTCTATGTACCTAAGCTGATCTCAAACTCTCTGG----- 231
QY 1678 TTGTTGTAATATTGTTCTTTTGTGAGGTAGGGCCCCAGACCAAAAAATAAATCT 1737
Db 232 -----GCTCAGGCGAGTCTCTCTGNTCCGGCTCCCAAGTGTGGGATTAATAG 281
QY 1738 TAGAATCAAAATCAGTGTGTTGGTTTGACCACTGTCACTTGAGAACCAAGTGTGACCG 1797
Db 282 GGGTGTGAGCCACACACCCAGCCCAATACTGTTTTTAAAAAAGCTTAAT----- 333
QY 1798 GGCCTCAGGAGTAGAGGTGATCTCTGCTCGAAGAGAAATAGAAATATTTCTCCGG 1857
Db 334 -----TGAGACATACTTCAATACCATACATATTTAAAGTGTCAAAATCAG 386
QY 1858 GCCAGGCTGTGGCTCATGCTGTAATCCAGCACTTTGGAGGCCCAAGGATGTGAT 1917
Db 387 TCCGGGCAATGGTGGCTCATGCTGTATCCAGCAATTTGGAAGGCCCGCAGGTGAT 446
QY 1918 CACCTGAGGTGAGGATTTCAAAACCAAGCCTGGCCAACTGGTGAACCCCTCTCTACTA 1977
```

```
|||||
Db 447 CACCTGAGCGCGGAGTTTCAGACCGGCTGCGCCAAACATGGTGAACCCCGTCTCTACTA 506
QY 1978 AAAATACAAAAAATAGCTAAGTGTGGTGGCGCATGCTGTAATCCACAGCTACTTGGGAG 2037
Db 507 AAAAT-CCAAAAATTAGCTGGTGTGATGGCTGTACCTGTGATCCCAAGTTACTCAGGAG 565
QY 2038 GGTGAGGCGAGGAGAAATTTCTTGAACCCGAGGCGCAGAGGTTGCAAGGAGAGATCACA 2097
Db 566 GCTGAGGCGAGGAGAAATCTCTTGGACCTGGGAGCGGAGGTTGCAAGTGTGAGATCACA 625
QY 2098 CCACGTGCACTCCAGCCTGGGAGAGAGCGAGACTTCTCTCAAAAAACAAAAACAA 2156
Db 626 CCACTTTACTCAAGCCTGGGTGACACAGCGAGATCTTGTCTCANAAAAAAGAAAAA 684

RESULT 15
AF101960 Human Homo sapiens genomic clone ptWB59.14, genomic survey
LOCUS AF101960 2971 bp DNA linear GSS 06-NOV-2000
DEFINITION AF101960 Human Homo sapiens genomic clone ptWB59.14, genomic survey
sequence.
ACCESSION AF101960
VERSION AF101960.1 GI:4193786
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2971)
AUTHORS Bepler,G., O'Brian,K.C., Kim,Y.C., Schreiber,G. and Pitterle,D.M.
TITLE A 1.4-Mb high-resolution physical map and config. of chromosome
segment 11p15.5 and genes in the LOH11A metastasis suppressor
region
JOURNAL Genomics 55 (2), 164-175 (1999)
MEDLINE 99134294
PUBMED 9933563
COMMENT Contact: Bepler G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
part of a 1.4 megabase contig including the LOH11A metastasis
suppressor region Bin T
Class: unknown.
FEATURES
source Location/Qualifiers
1. .2971
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="11p15.5"
/clone="ptWB59.14"
/clone_lib="Human"
ORIGIN
Query Match 6.5%; Score 255.8; DB 28; Length 2971;
Best Local Similarity 65.4%; Pred. No. 5e-30;
Matches 478; Conservative 0; Mismatches 177; Indels 76; Gaps 4;
QY 1442 TTTTATTTTATTTTGGAGAAAGAGTCTGTCTTTTCCACCTGGCTGGAGCGCAGTGTGT 1501
Db 478 TTTATTTATTTTCCGAGACAGTGTCTCACTCTGTGTTGCCAGACTGGAGTGCAGTGGCAC 537
QY 1502 GATCTCAGCTCAGTCGCACTTTCGCTCCCGGGTTCAGCAATTCCTGCTTCAGCCTC 1561
Db 538 AATCTTGGCTCAGTCGCAACCTCCGCTCTCCCTGGGTTCAAGCGATTCTCTGCTCAGCCTC 597
QY 1562 CCAAGTAGCTGGGACTACAGGCACTTCCACCATGCGCCAGCTGATTTTGTATTTTACT 1621
Db 598 CCAAGTAGCTGGGATACAGGCGCCCGCCACCAACAGCTAATTTTGTATTTTACT 657
QY 1622 AGAGATGGGATTTTCACTTTTGTGGCCAAAGCTGGTCTCAAACTTTTGTCTGCATATAATGT 1681
Db 658 AGAGATGAGGTTTCAACAAGTTGGCCAGGCTGTCTCAAACTCTT----- 702
```

```
QY 1682 TGTAACTATTGTTCTTTTGTCTGAGGTAGGGCCCGCCAGACCAAAAAAATAAATCTTTAGA 1741
Db 703 -----GACCTCGTGATCCACCCACTCAGCCTCCCA 733
QY 1742 ATCCAAATACAGTGTGTGGTTTGACCACTGTCTTGTGAGAACCAACAGTGTGACCCAGGCGC 1801
Db 734 AAGTGTCTGGGATTACAGGCATAAGCCGCACTCTGGCCCAACATTATAT----- 783
QY 1802 TCAGGAGTAGAGGTGATCTCTGCTCGAAAGAGAAATAGAATGAAAAATATTCTCCGGGCGCA 1861
Db 784 -----TTTATGGTGAAGAAGACTGAAATGTTTCTTAAGATCAGAAAAACAGCCA 827
QY 1862 GGCCTGTGTGGTCAATGCTGTAATCCACAGCACTTTGGAGGCCCAAGGCATCTGGATCACC 1921
Db 828 GGCACAGTGGCTCACGCTGTGTAATCTTGGCACTTTGGAGGCCCGAGCGGCGCAGATCAC- 886
QY 1922 TGAGGTGAGGAGTTCAAAACCCAGCCTGGCCAAACATGTTGAAACCCCGTCTCTACTAAAAA 1981
Db 887 -GAGGTGAGGAGATCGAGACCATCTCTGGCTAAACACGCTGAAACCCCGTCTCTACTAAAAA 945
QY 1982 T----ACAAAAAATTAGCTAAGTGTGTGGCGCATGCTGTAATCCACAGCTACTTGGGAG 2037
Db 946 TAAAAATAAAAAATTAGCTGGTGCAGTGGTGGCACCTGTAGTCCAGCTACTTCGGGAG 1005
QY 2038 GGTGAGGCGAGAGAAATTTTGAACCCGAGGCGCAGAGTTTGCAGTGAAGCGAGATCACA 2097
Db 1006 GCTGAGGCGGAGAGATGGTGTGTAACCCAGAGGCGGAGCTTGCAGTGAGCCAGATCATG 1065
QY 2098 CCACTGCACTCCAGCCTGGGCGGAGAGAGAGACTTCTCTCAAAAAACAAAAACAAA 2157
Db 1066 CCACTGCACTCCAGCCTTAGCGCACAGGGCGAGACTCCGACTCAAAAAAAGAGATTA 1125
QY 2158 AGAATTAAAGCA 2168
Db 1126 GAAATGAACA 1136
```

Search completed: May 13, 2004, 17:08:27
Job time : 14558.7 secs